DOCUMENT RESUME

ED 234 495 EA 016 082

AUTHOR Sanders, Susan; And Others
TITLE The Computer in Educational Decision Making. An Introduction and Guide for School Administrators.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO ISBN-0-89466-281-3
PUB DATE 78
NOTE 174p.; Published by Time Share Corporation, Hanover, NH.
PUB TYPE Guides - Non-Classroom Use (055) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Administrators; *Computer Oriented Programs; *Computer Programs; *Decision Making; *Educational Administration; Educational Planning; Enrollment Projections; Evaluation Methods; Guidelines; Linear Programing; Management Systems; Mathematical Models; Models; *Operations Research; Planning; *Problem Solving; Simulation

ABSTRACT

This text provides educational administrators with a working knowledge of the problem-solving techniques of PERT (planning, evaluation, and review technique), Linear Programming, Queueing Theory, and Simulation. The text includes an introduction to decision-making and operations research, four chapters consisting of indepth explanations of each technique, and instructions on the use of computer programs. PERT is used for planning and analyzing stages or activities in project development. The chapter includes step-by-step instruction in its use--with illustrated charts and examples--and an explanation of the GCPATH program for a PERT computer analysis. The third chapter centers on the use of Linear Programming, a mathematical technique designed to solve mathematically stated problems. It contains examples of the computer program LPRG and illustrates its use in several different situations. Queueing Theory, the subject of the fourth chapter, is a method for analyzing waiting line problems. Presented are the basic elements of the theory, discussion of its use, and an outline of solutions to waiting problems with the QUEUE computer program. The final chapter introduces Simulation Technique, where a variety of simulated solutions to problems are used instead of real life situations. Practical examples illustrate the application of the technique and the use of the computer programs ENROLL, SUBST, and BUSRTE. (MD)

THE COMPUTER IN EDUCATIONAL DECISION MAKING

An Introduction and Guide for School Administrators

Susan Sanders
Stuart Speedie
Duane Richardson
Ralph Van Dusseldorp
Antoinette Ellis
Hollis Adams

Foreword by
Richard Otte
Project Officer
Technological Applications Division
National Institute of Education

Director, Computer Technology Program
Judith B. Edwards

Northwest Regional Educational Laboratory
710 S.W. Second Avenue
Portland, Oregon 97204

TimeShare
A Houghton Mifflin Company
Hanover, New Hampshire 03766

This work was developed under a grant and contracts with the National Institute of Education, U.S. Department of Health, Education and Welfare. However, the content does not necessarily reflect the position or policy of that Agency, and no official endorsement of these materials should be inferred.

ISBN 0-89466-281-3

CONTENTS

I. AN INTRODUCTION TO EDUCATIONAL DECISION MAKING AND OPERATIONAL RESEARCH

II. PERT: A PLANNING AND ANALYSIS TOOL

III. LINEAR PROGRAMMING

IV. QUEUEING THEORY

V. COMPUTER SIMULATION

Foreword

Among modern planning and decision making tools, the several techniques developed in operations research in the 1940's stand as some of the most powerful and versatile tools for modern management. Until recently, the complexity of computations underlying these techniques precluded their extensive use in educational administration. However, as computer computational support is becoming increasingly—indeed commonly—available to educators, and the necessary computer programs are now readily obtained, operations research techniques are beginning to make their unique contribution to the modern educational administrator's battery of planning and decision making tools.

The purpose of this book is to provide the educational administrator with a working understanding of the most useful operations research techniques and experience in using computers to provide the background computations required by each. The emphasis throughout the text is on the practical application of the techniques to educational problems. The text presupposes only that the user has access to a computer. The use of computer programs and the interpretation of their results for decision making are taught in the text, and all complex mathematical manipulations underlying the application of the techniques are left to the computer.

This book provides a practical, non-technical introduction to operations research techniques in educational planning and decision making and as such should be a valuable teaching vehicle for practicing school administrators as well as for students and professors of educational and business administration.

Development of this book by the Northwest Regional Educational Laboratory was sponsored by the National Institute of Education (NIE), Washington, D.C. NIE has demonstrated a continuing commitment to the sponsoring of research which results in the improvement of education through the application of modern technology. NIE is particularly pleased to have sponsored this project which facilitates the application of a powerful technological tool (the computer) and a powerful set of management tools (operations research) to the improvement of education. This product fills an especially vital need today as strains on school budgets together with the public call for improved management and accountability demand more efficient, effective management in the complex and expanding enterprise of education.

Richard B. Otte
Education Technology Specialist
National Institute of Education

1 An Introduction to Educational Decision Making and Operations Research

ABOUT THE BOOK

The purpose of this text is to provide the practicing and future educational administrator with a working knowledge of some of the most useful tools of modern management—the operations research techniques of PERT, linear programming, queueing theory, and simulation—using both hand and computer calculations. Throughout the text, practical exercises are given which require the reader to apply the information just studied. Answers are included in Appendix II. An understanding of basic algebra will equip the reader to handle all hand calculations required; the text itself guides the user through the simple steps of running the numerous computer programs, assuming only that the programs have been loaded onto the available computer and that the user can obtain any needed guidance for getting access to the program from the terminal.

This first chapter is a basic introduction to the four operational research techniques dealt with in the text and briefly discusses their practical applications in the educational setting. The chapter can be used effectively as an advance organizer for the subsequent four chapters, which explain in detail the use of the several techniques as aids in complex decision making.

COMPLEX DECISION MAKING IN EDUCATION

The primary focus of an educational administrator's job is on decision making, and the scope of the administrator's decision-making responsibilities extends across the entire educational enterprise. In many day-to-day educational problem situations, alternate solutions are so clear and well-defined that deciding on the best solution can be intuitive. Most routing problems, such as which teachers to assign to lunch duty or which of two new teachers to hire, are regularly solved by familiar, semi-intuitive procedures. An increasing number of educational problems confronting administrators are, however, so complex and involve so many possible solutions that it is difficult or impossible to pick out intuitively the best solution.

One example of the large and complex problems that educational managers face is that of minimizing costs in all phases of education while maintaining or improving quality. For example, the administrator must typically make the decisions necessary to hold down the costs of the school lunch program while at the same time assuring that each child receives a nutritionally adequate meal. He or she must minimize the cost of buying new textbooks while making sure that the quality of instruction remains high. He or she must make the decisions required to negotiate teaching contracts while simultaneously holding the line on costs and ensuring adequate or improved instruction; bus schedules must be constructed so that all children are transported efficiently at a minimum cost; building maintenance must be scheduled to minimize both cost and interference with instruction. Clearly, arriving at solutions to problems requiring simultaneous minimizing and maximizing of effects must involve more than intuition.

Planning is another area in which large and complicated problems arise. Here the administrator faces such questions as: How can a curriculum revision be carried out so that all the materials will be ready by the start of school next year? How do I make sure that all the small tasks necessary to complete this revision get done in the right order and with a minimum of time wasted? What is the best sequence of activities for carrying out a revision of the administrative structure? What tasks have to be completed and in what sequence must they be done in order for me to be ready for upcoming salary negotiations? What is the best and most efficient way to carry out a needs-assessment study for my district? What are the activities and what is the sequence for preparing my school's budget for next year? The solutions to such planning problems are generally beyond the scope of simple intuitive decision making.

In addition, the educational manager is often faced with problems of prediction. He or she must have some rational basis for answering questions like: How is the population in this district going to change over the next ten years? How many and what kinds of physical plant facilities will the district need in the future? How must the teaching staff change to meet expected enrollment changes? What new equipment must be purchased in order to meet the future needs of the students, faculty and community?

For each of these complex problems, the number of possible solutions is large, but there is usually only a small number of solutions that will meet all criteria for a successful solution. That is, there are only a few solutions which will simultaneously minimize costs and maximize some other important variables (such as achievement), which will create the most efficient plan, or which will predict in the optimum manner. These types of problems also require more than intuition.

The administrator's situation is made additionally challenging by the variety of problems to be solved, including such concerns as racial integration as mandated by the Supreme Court, accountability, static or decreasing monetary support in the face of steadily rising costs, changing staffing patterns, and demands for increased quality of educational practices.

In all these educational administration problems, two things are required of the administrator: he or she must make decisions to solve the problem, and must have tools beyond insight and intuition with which to work out the decisions. We will first explore exactly what is involved in the decisions themselves, and then consider techniques for optimizing the effectiveness of the decision process.

To perform the job of increasingly complex decision making with maximum effectiveness, present-day administrators have found they can make excellent use of all modern management tools they can command. Among the newest such tools are operations research techniques.

OVERVIEW OF OPERATIONS RESEARCH TECHNIQUES

The need for special research into large operations was first recognized as critically important in the large development and production projects of the Defense Department and major industries during the 1940s and 1950s. Here such problems as the development of effective complex deployment strategies, the design and production of new weapons systems, and the production of huge airliners were first encountered.

In order to find the best possible solutions for these problems, several scientifically based, highly technical methods were developed. The applications of these techniques over the years became known as operations research. Although these techniques were originally designed for specific projects, they have been generalized to the point where they have become widely applicable tools for modern management. They provide a systematic way of looking at problems and of generating data that will provide a base on which to make effective decisions.

To the degree that one faces complex multifaceted problems, the educational administrator will find some of these techniques important aids, perhaps better than methods previously available. The most useful operations research techniques for modern educational administrators are:

PERT
Linear Programming
Queueing Theory
Simulation

All four are presented in depth in Chapters 2 through 5. As an introduction to these decision-making tools, a brief description of each follows.

PERT

The acronym PERT stands for *P*lanning, *E*valuation and *R*eview *T*echnique. PERT is a highly effective procedure for organizing information about multistage projects. The technique is used primarily for planning and analyzing project stages or series of activities which are oriented toward some goal and are interdependent. The key feature of the PERT technique is the PERT chart, which graphically displays the activities involved in a project and their interconnection. The PERT chart forms the basis for detailed analysis of time elements, critical sequences of activities in the progress of the project, and scheduling requirements.

An example of a PERT chart is shown in Figure 1-1 for the simple project of preparing for a speech. The arrows represent the sequential activities involved in this project.

To see how PERT might be used in a typical educational problem situation, suppose it has been decided that your district administrative structure is to be reorganized. Obviously, the problem is to complete the reorganization successfully, but a great many different, interrelated activities with varying and variable time constraints will be involved. To approach this problem with the tool of PERT charting and analysis, the first step to be taken is the determination of all activities which must be a part of the reorganization project. Using the PERT chart procedure, the activities are laid out in diagrammatic form to show the order with respect to one another and to the final goal—the completed reorganization. The construction of the chart systematically clarifies crucial predecessor activities and interconnection points. Once the chart is complete, the time required for each activity can be entered on the diagram, and further analyses can be carried out to identify where time factors must be manipulated to ensure the timely flow of activities, where slack time may be available, and what sequence of activities involves the least slack or no slack and hence needs special attention. On the basis of the completed time analysis, the scheduling of each activity in the project can take place.

In summary, PERT is a technique that can be of immense help to administrators in planning the successful and timely completion of complex projects.

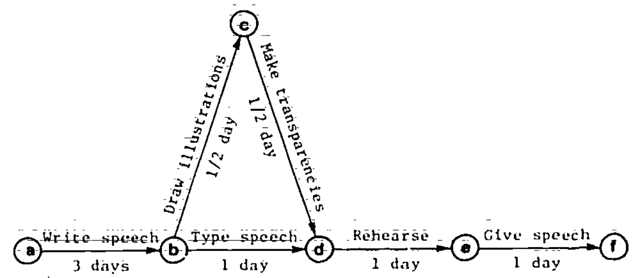


Figure 1-1.
Simple PERT Chart. This chart may be interpreted as follows. Writing the speech must take place before either the illustrations are drawn or the speech is typed, and the writing will take three days. Drawing the illustrations must be done before the transparencies are made, and the drawing will take ½ day. Both typing the speech (1 day) and making the transparencies (½ day) must take place before the speech is rehearsed. Finally, the speech must be rehearsed (1 day) before the speech is given (1 day).

LINEAR PROGRAMMING

Linear programming is a mathematical technique designed to solve problems that can be stated in terms of constraints and a single measurable goal. Linear programming is applicable to problems in which the constraints and goal can be stated mathematically. The purpose of linear programming is to find *the largest or smallest possible value of the goal that conforms to all of the restraints.*

Let's consider a problem that may be solved by the method of linear programming. Suppose that some of the teachers in your school are having trouble deciding how much time they should spend in large group and individual instruction. The need, as you see it, is to decide how to divide teacher time between large group and individual instruction. Further, suppose that the average class size for your school is 30 students and that each teacher spends five hours (or 200 minutes) each day in group and individual instruction. This constraint can be expressed mathematically: the sum of the time spent in group instruction plus 30 times the amount of time in individual instruction (since there are 30 students who must be individually instructed) must be equal to 300. The second and third constraints are that the teacher must spend some time in each of the two modes of instruction; the amount of time spent in group instruction must be greater than zero, and the time spent in individual instruction must also be greater than zero. Mathematically we can express these constraints as follows:

Constraint 1: Group Time + (30 × Individual Time) = 300
Constraint 2: Group time > 0
Constraint 3: Individual Time > 0

Finally, we must formulate the goal as a mathematical expression. The goal is for each teacher to spend time during a school day in such a manner that the students will derive the most benefit from the instruction they receive. Before we can formulate the goal, we must decide on and quantify the value of group instruction relative to the value of individual instruction. For purposes of using linear programming to solve this problem, specifying individual instruction simply as more valuable than group instruction is not sufficient. We must quantify how much more valuable individual instruction is than group instruction. Suppose we decide that individual instruction is five times more valuable than group instruction. Now we can express the goal mathematically by:

benefit to students = Group Time + (5 × Individual Time)

With these quantified expressions, the techniques of linear programming can now be used to calculate a single optimum combination of group instruction time and individual instruction time. This optimum combination of times would make the value of the goal (benefit to students) as large as possible subject to the imposed constraints. Linear programming is a powerful decision strategy when the solution to a problem involves minimizing or maximizing a goal expressible in mathematical terms, within the limits of constraints that can also be expressed mathematically.

QUEUEING THEORY

Queueing theory is a method for analyzing waiting line or queue problems—that is, problems in which there are one or more service "facilities" and where units of some kind require the services of these "facilities" at various times. Service facilities may include school lunch servers, film projectors, ditto machines, district maintenance departments, guidance counselors, examinations, learning packages, or teachers. The units of the queue or waiting line may be students, teachers, committees, typewriters, automobiles, school buildings, and so forth. These units (called "customers") do not have to be physically waiting in line, but they must require service by the facility.

The educational setting abounds in such waiting-line situations, often involving large numbers of customers and complex questions concerning needs, time, and available funds. Queueing theory provides a mathematical framework within which to study the key aspects of the queueing situation and to determine such things as what the average length of the waiting line is, how long a customer just arriving at the waiting line can expect to wait, and how much of the time the service facility is idle. Corresponding information can be systematically generated for hypothetical alternative waiting-line situations in such a way that practical comparisons can easily be made.

To gain a clearer introductory view of what queueing theory provides, consider the following example. Suppose your district has two major building repair crews. Assume you have just been charged with deciding whether to add a third crew. The goal here is to determine whether two or three repair crews are needed. Some of the practical constraints to be taken into consideration might be such things as the amount of money available for repairmen or the amount of time crews may be idle. Approaching this problem with queueing theory methods, the repair crews are identified as the service facilities and the customers are the school buildings needing repairs. If you are able to determine the average time it takes a crew to repair a school building, you can use queueing theory to generate the information you will need to solve the problem. Among other data, this technique can provide you with information about the average waiting time for repair, the average number of buildings waiting to be repaired with two or three service crews, and the probabilities of the crews being idle. In a case like this, you would have to make the final decision based on the criteria and constraints which apply. For example, your final decision on whether to have two or three crews might depend on the criterion that no building should wait over two weeks for major repairs.

This example is one of many possible situations in education for which queueing theory provides a powerful decision-making aid.

SIMULATION

Simulation is an operations research technique for formulating and operating models of problems. It is a technique for translating the problem situation into a working model of the problem. When we speak of "model" in this context, we are referring to a representation of the real-world situation which is simpler and easier to understand and manipulate than the real-problem situation. For instance, a mathematical equation may be a model of a real system; a set of logical steps can be the model of a decision process. The principal value of a model to the administrator is that he or she can try out a number of different solutions to the problem on the model without taking the time or money that would be required if these solutions were tried out in the real-decision context.

It should be noted that PERT, linear programming, and queueing theory are *specialized* types of simulation. PERT uses an analysis of a project to construct a diagrammatic model of the project. Linear programming develops mathematical equations that are a model of the problem situation. Queueing theory develops essentially a mathematical model of problem situations where there are customers waiting to use a service.

The simulation technique, in general, is used to analyze the problem situation into its components and to put these components into a simplified system representing that situation. One then develops a number of possible solutions to the problem. These solutions are applied to the simplified system, and the outcomes or results of the various solutions are observed in the model. Thus simulation is primarily a tool for assessing the impact of various possible decisions on the problem situation.

Let's take an example to see how simulation works. Suppose that you are interested in determining the need for new school buildings in your district over the next ten years. The basic problem may be stated in the question, "How is the population of children in my district going to change over the next ten years?" In consultation with an expert in the area of demographic change, you would analyze the situation and identify the constraints and controllable variables. Using relationships among these constraints and controllable variables as specified by a theory of demographic change, you and the expert would develop a model of demographic change for your district.

This model would relate changes in the school-age population to such factors as economic growth rate, birth rate, age-range distributions in the population, number of housing starts, changes in the geographic distributions in the population, and rate of industrial development. In order to find solutions to the problems implied in the situation, you would try various combinations of the input factors. The model would then operate on these input factors to yield an outcome in terms of changes in school-age population. Your decision on how many school buildings your district needs in the next ten years would be a function of the input data you consider the most probable for your area.

As the above example suggests, simulation is an effective technique for investigating the problem situations and providing information as to the ramifications of many alternate decisions.

These, then, are four major techniques in operations research—PERT, linear programming, queueing theory and simulation. In the following chapters you will get a better understanding of these techniques and how they can be applied to the problems of the educational administrator.

OPERATIONS RESEARCH USING THE COMPUTER

In complex problem situations, the operations research techniques discussed above will prove to be excellent ways of organizing information and generating data for decision making, but the calculations required can be tedious and time-consuming for the heavily scheduled administrator. Since these techniques have found a broad range of uses in management today, however, computer programs designed to perform the calculations have become readily available.

Educational administrators will often find that the manager of the computer center in their district, or whose services they use, can obtain programs that perform PERT, queueing, or linear programming either from the library of programs available with their system or from a university computer center or user's group. Many commonly used simulation programs are also available from such sources. In most cases the input information needed to run the programs will be the data regularly required by the technique, and output from the computer will normally be easily interpreted by the administrator familiar with the technique involved. The running of programs will, in most instances, be left up to the computer center personnel.

At the end of each of the following four chapters, after each technique is described and hand and computer calculations are both explained and demonstrated, a final discussion is given concerning how to obtain and use programs for that specific technique.

BIBLIOGRAPHY AND SUGGESTED REFERENCES

Optner, Stanford L. *Systems Analysis for Business Management*. Third Edition. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.

Shamblin, James E. and G. T. Stevens, Jr. *Operations Research: A Fundamental Approach*. New York: McGraw-Hill, Inc., 1974.

Siemens, Nicolai, C. H. Marting, and Frank Greenwood. *Operations Research*. New York: The Free Press, 1973.

Van Dusseldorp, Ralph A., D. E. Richardson and W. J. Foley, *Educational Decision-Making Through Operations Research*. Boston: Allyn and Bacon, 1971.

Wagner, Harvey M. *Principles of Operations Research*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.

2 PERT: A Planning and Analysis Tool

CHAPTER PREVIEW

The PERT planning and analysis tool is a highly effective technique for structuring and analyzing milistep projects. This chapter is designed to give the reader a working knowledge of PERT organizing method, including the PERT network, and a clear understanding of the most important uses of PERT in administrative project planning.

Experience is provided in using computer programs to carry out PERT analyses of complex projects typical of the ones educational administrators frequently must plan.

Throughout the discussion of PERT methods, practical but simple examples are used to illustrate the application of PERT in educational settings, although it must be kept in mind that PERT has its greatest use and benefits in more complex project planning. At key points, REVIEWS are interspersed which reiterate definitions of new terms and provide progressive exercises on the methods previously covered. A final exercise requires the reader to use all the methods involved in PERT network and analysis on a typical project using a computer program.

CHAPTER AIMS

After completing this unit, the reader will be able to command the basic PERT analysis and planning methods and to use them to solve problems involving complex scheduling of activities and variable time lines. Mastery of the material will also provide the reader with the skill to use PERT analyses to communicate with others about projects and to use computer programs to perform complex PERT analyses as needed in the actual administrative setting.

INTRODUCTION TO PERT

A number of large-scale research and development efforts has been initiated by the Government and private industry during the last two decades. One such project was the Polaris ballistic missile. At the time of this project's inception, the designers needed to plan and execute the project as quickly as possible; however, a number of unanswered questions existed: what research was needed and how was it to be planned? What was the length of time these research projects would take to complete? What were the stages of development and testing which the successful completion of the project would entail? At what date could the United States expect to have an operational Polaris ballistic missile? To answer these questions the Navy developed a project planning system called PERT, for Planning, Evaluation, and Review Technique.

PERT is a planning and scheduling technique which provides for systematic analysis and planning of projects in which there is a degree of uncertainty. It is valuable when the exact time needed to complete a project is not precisely known but can be estimated. Fundamental to the PERT technique is the development of a graphic diagram or network of the sequence of all activities involved in a project, from its inception to its completion. Thus PERT can be simply described as a tool for planning and analyzing a project involving a series of activities when the duration of the activities is uncertain but capable of being estimated.

A key feature of the PERT network and analysis is the method of finding the critical path, or sequence, of activities in the project network. For the manager of any project, this amounts to determining in advance the sequence of strategic activities which involves the least room for delay and which may therefore require special attention either before or during the project.

PERT has four particular advantages for the decision maker. First and foremost, it is a powerful tool with which to structure thinking about many of the tasks confronting the decision maker. Second, it offers a versatile, systematic methodology for the analysis and planning of multistage projects. Third, it provides the decision maker with a useful device for communicating project plans to all levels of staff and management. Finally, PERT affords the manager a practical basis for effective time management throughout the most highly complex and lengthy projects.

PERT IN EDUCATIONAL ADMINISTRATION

To the extent that educational administrators are responsible for the management of time and manpower, PERT is an enormously useful tool. The following tasks are just a few of the frequently encountered administrative projects for which PERT is appropriate and useful:

1. Budget preparation
2. Analyzing procedures for the purchase of materials and supplies
3. Planning a feasibility study of administrative organization
4. Planning and controlling an educational needs survey
5. Analyzing scheduling procedures
6. Planning a curriculum evaluation
7. Analyzing media scheduling
8. Planning and analyzing maintenance procedures

Each of the projects above involves multiple activities with complex timetables governing their completion and interconnections. Without a methodology such as PERT for organizing and scheduling complex projects, the administrator may lose considerable time developing organizational schemes for structuring and implementing projects with multiple activities. With PERT the framework for organizing factors in the project is already at hand, and the central work of structuring and analysis can begin without delay.

PERT ANALYSIS: FOUR STEPS

PERT assumes that the work under consideration constitutes a project. Cook defines a project as "an organization unit dedicated to the attainment of a goal—generally the successful completion of a development product on time, within budget, and in conformance with predetermined performance specifications."[1]

[1] D. L. Cook, *Better Project Planning and Control Through the Use of System Analysis and Management Techniques* (1967), ERIC ED019 729.

Thus projects are finite in that they must have a definite beginning and ending. They must be complex in that they require a mix of human and material resources applied tc a series of related activities. They must be homogeneous, in that they must be capable of being distinguished from other projects. And they must consist of a well defined collection of interrelated activities.

Given such a project, the following are the steps involved in a PERT analysis:

Identify the activities involved in the project.
Construct a network by pictorially representing the relationships between activities.
Determine the time necessary to complete each activity.
Compute the time required to complete the entire project, the critical path of activities through the project, and the times at which each activity must be completed in order for the entire project to be completed on time.

We will consider each of these steps in turn.

I. IDENTIFY THE ACTIVITIES

The first step in PERT analysis of a project is to identify the activities which make up the project. That is, the project must be broken down into its component activities. The typical project using PERT methods in its planning might have anywhere from 25 to 100 activities on the list. For illustrative purposes in our text, however, we will use simplified examples. As the first, suppose you want to identify the activities involved in presenting a new curriculum package to your school board. Assume that you were an *ex officio* member of the curriculum revision committee and were therefore not completely aware of the proceedings. Probably the first thing that you would do is to get a copy of the committee's report and a sample of the curriculum materials. Then you would have to prepare your talk—that is, outline, write, and type your presentation. You might also want some overhead transparencies. Finally, you must complete the activity of giving the talk. The list below gives the major activities:

Get curriculum committee report.
Get curriculum materials samples.
Outline presentation.
Write presentation.
Type presentation.
Make overhead transparencies.
Make presentation to board.

Because of the simplicity of this example, we have listed the activities in the sequence in which they may occur, although that is not necessary. Lists may often reflect a very random order.

Notice that in the list above, the activities comprise a clearly defined set of interrelated activities, each of which is discrete and has distinct beginning and ending points. Recall that for PERT to be applicable, a project's activities must exhibit these characteristics. The above list, then, confirms that the project may be analyzed using PERT. This is not to say, however, that it is the particular list you would have constructed for this project. Your list could have included quite different activities, each of which had discrete and independent starting and ending points. The important thing is that PERT is an administrative tool, its use by any individual will necessarily be slightly different. *The correct list for any project is the one that the particular administrator feels is appropriate for his or her purposes.*

REVIEW

Terminology

1. PERT (*P*lanning, *E*valuation and *R*eview *T*echnique): a tool for planning and analyzing projects with a definite beginning and end and involving well defined activities that can be ordered in time.

Exercise

1. List the essential activities that you think are involved in setting up an elementary school budget.
2. List the activities you think are involved in conducting year-end staff evaluations.

II. CONSTRUCT A PERT NETWORK

The second step in a PERT analysis is to construct a PERT network. This network is a pictorial representation of the interrelationships among the activities comprising the project.

To illustrate the construction of a network, let's take a simplified version of the board presentation project in which the activities are identified as:

Make presentation to the board.
Outline the presentation.
Write the presentation.

To keep track of all the activities in a project, it is conventional to make a complete list of all activities involved (not necessarily in any order) and to identify each by an identification number (again, not necessarily in any order). A typical activities list for our sample project is shown in Figure 2-1.

ID	*Activity Description*
1	Make presentation to the board
2	Outline the presentation
3	Write the presentation

Figure 2-1.
Sample activities list.

As you can see, the list gives the verbal description of each activity in the center (other information will be added to the right later on) and places the identification number in the left-hand column. The identification numbers give a clear and convenient way to refer to each distinct activity in the construction of the network.

Event: Before we construct the network, the concept of an *event* must be introduced. An *event* marks the beginning or the end of an activity. An event, as distinguished from an activity, has no time duration. The initiating event represents the start of a particular activity and the terminal event represents the point at which the activity is completed. Thus, in our simple example, the first event would mark the start of the activity of outlining the presentation. The first activity is terminated by the ending event. At this point, the activity which follows the outlining of the presentation is initiated, that of writing the presentation. Since it directly follows the ending event of Activity 1, its starting event is concurrent with the ending event of Activity 1, and it terminates by its own ending event. And so on.

We can now construct a PERT network for our simple example, showing each activity, in order, and showing their starting and ending events.

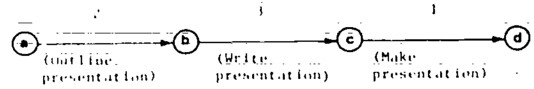


Figure 2-2.
PERT Network for Board Presentation Project.

You can see that the diagram in Figure 2-2 represents the project in terms of the flow of activities and events. The arrows represent activities, with the identification number of the activity on the top of the arrow. Thus —2→ stands for Activity 2, which, in our example, is outlining the presentation. A lettered circle denotes an event; intermediate events may represent both the termination of one activity and the initiation of another activity. Note that the diagram flows from left to right, and specifies which activities must be completed before other activities may begin.

The network in Figure 2-2 shows diagrammatically that in the simplified board presentation project, the first event (event *a*) is the beginning of Activity 2 (outlining the presentation). This is followed by the activity of outlining the presentation and the termination of the activity (event *b*). This, in turn, is followed by Activity 2 (making the presentation to the board), which ends in its terminating event (event *d*). Thus the network is a visual representation of the order in which activities must take place for the project to be completed.

Now let's return to the complete board presentation project for which we identified the seven activities on page 9:

Get curriculum committee report.
Get curriculum materials samples.
Outline presentation.
Write presentation.
Type presentation.
Make overhead transparencies.
Make presentation to board.

Before beginning to construct the network for this project, we will make a conventional activities list and assign an identification number to each activity for our future reference. Our list might look like that in Figure 2-3.

ID	*Activity Description*
1	Get curriculum committee report
2	Get sample curriculum materials
3	Outline presentation
4	Write presentation
5	Type presentation
6	Make overhead transparencies
7	Give presentation

Figure 2-3
Activities list for complete board presentation project.

Remember that activities may be listed in any order and that numbers may also be assigned in any order that is convenient. Using this list, one can construct the activity network.

As is conventional, we construct the network from left to right, beginning with the first activity. An arrow is drawn for each activity, in logical order, by marking its start and end with a lettered circle representing the starting and ending event. Since PERT assumes a continuous sequence of interconnected activities, the lettered circle marking the end event is used simultaneously as the starting event of the next activity directly following it. For example, —3→ⓔ—4→ⓓ—5—. The one exception is obviously the final activity of a project—for example, —7→ⓕ—8→ⓖ. Also note that the initial starting event naturally has no predecessor activity, hence is marked by a lettered circle which has no arrow leading to it. For example, ⓐ—1→ⓑ. When a project has two or more initial activities that start independently, each is represented by a separate arrow beginning with its own starting event.

With these conventions in mind, we can now construct the network for the board presentation project. From the activities list, we see that we have two initial events without predecessors—Activity 1 (get curriculum committee report) and Activity 2 (get sample curriculum materials); these will be represented by two separate activity arrows each of which has its own independent starting event. But, what about their ending event? We know that, except for the final ending event of a project, the ending event of each activity marks the beginning event of the next activity. What activity follows getting the curriculum report and what activity follows getting sample materials? Obviously, Activity 3 (outline the presentation) follows both Activities 1 and 2—that is, the beginning event of Activity 3 is the ending event for both Activities 1 and 2. Therefore, the network must begin like this:

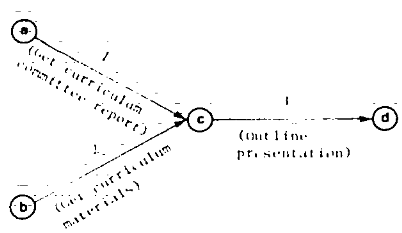


Figure 2-4.
First Three Activities in Board Presentation Project.

Notice that each event is lettered and that event *d* marks the end of Activity 3.

To continue the network, we need to determine what activity follows the ending event (outlining the presentation)—that is, what activity Activity 3 immediately precedes. Since Activity 4, writing the presentation, can begin immediately after and not before the end of Activity 3, we can draw this activity in and produce Figure 2-5.

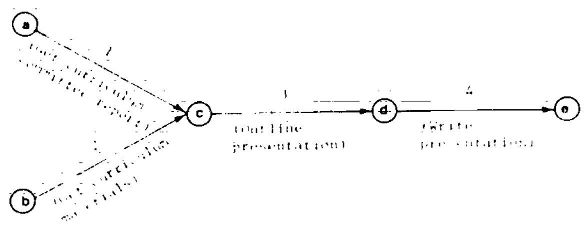


Figure 2-5.
Activities 1 to 4 of Board Presentation Network.

What activity immediately follows and cannot begin before the end of writing the presentation? Clearly, it is typing the presentation, Activity 5.

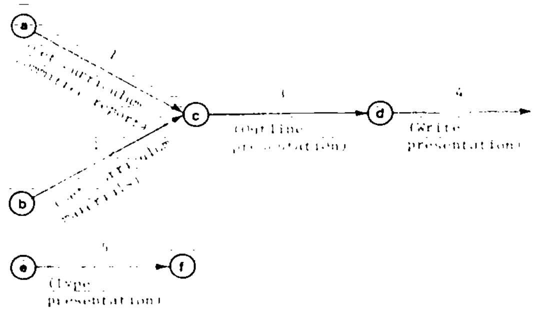


Figure 2-6.
Activities 1 to 5 of Board Presentation Network.

Now, what follows Activity 5, (typing the presentation) and cannot begin before it? Activity 6, making overhead transparencies? But these could be made before the presentation is typed. They could be made even before the presentation is written, as Activity 4. But they could in all likelihood not be made before the presentation is outlined, Activity 3. Therefore, the immediate predecessor of Activity 6 is Activity 3.

The only activity left which could follow Activity 5 (typing the presentation) is Activity 7, making the presentation. This Activity can be added to the network, and as it is clearly the terminating activity of the project, its ending event *g* is the termination point of the project.

We have now only Activity 6, making overhead transparencies, to enter in the network. As we discovered above, Activity 6 is immediately preceded by Activity 3, the outlining of the presentation. So, we could enter the arrow for Activity 6 showing its starting event to be concurrent with the ending event of Activity 3, event *d*.

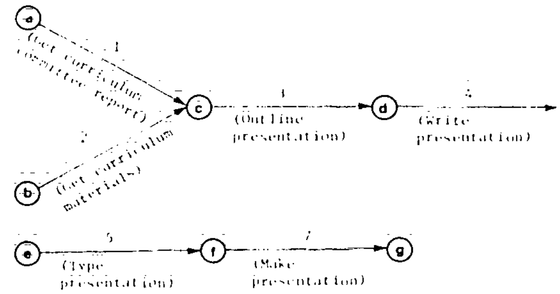


Figure 2-7.
Activities 1 to 7 of Board Presentation Project.

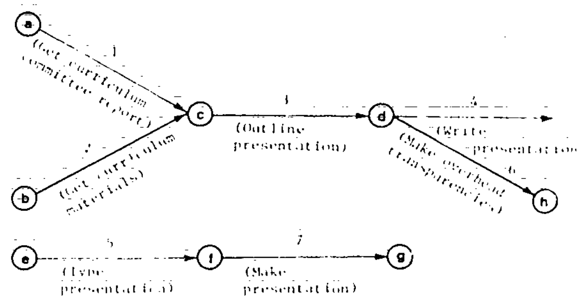


Figure 2-8.
Activity 6 following Activity 3 in Board Presentation Network.

At this point we have not yet finished constructing the network. As you will notice, Activity 6 appears to end in midair at event *h* and is not shown to precede Activity 7 (making the presentation). We must reconstruct the network in such a way as to show that Activity 7 is immediately preceded by both Activity 5 (typing the presentation), and Activity 6 making transparencies).

Since PERT conventions do not allow curved arrows, we cannot curve the arrow for Activity 6 to run from the ending event *d* to the starting event *f*. However, we can place Activity 6 on a straight arrow between the ending event *d* and the starting event *f*; then we can draw the arrow for Activity 4 (writing the presentation) starting at event *d* and proceeding above the main line of arrows to end in event *e*. We can draw a descending arrow for Activity 5 (typing the presentation) between its starting event *e* and at its ending event *f*, as shown in Figure 2-9.

Now, reading Figure 2-9 from left to right, you can see that the network details the order of activities in preparing the presentation to the school board and indicates exactly which activi-

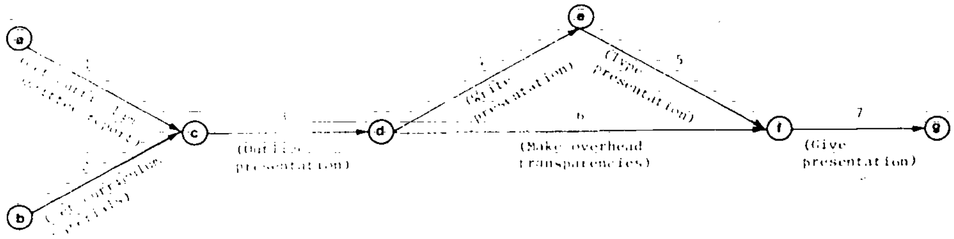


Figure 2-9.
Complete Network for Board Presentation Project.

ties must precede other activities so that the project can be successfully completed. In constructing networks, administrators will often find that several tries are required before the correct sequence can ultimately be diagrammed, showing the exact concurrence of ending and starting events from the beginning of the project to the finish.

If the administrator stops exploring PERT even at this point, a valuable method has clearly been acquired for organizing projects. However, there are more elements of PERT analysis which can be used in planning and implementing projects.

Once the network has been adequately constructed, the administrator can add the correct immediate predecessor for each activity to the activities list for future reference and for later use with computer programs. As an example, Figure 2-10 illustrates the activities list for the board presentation project with immediate predecessors filled in.

ID	*Activity Description*	*Immediate Predecessors*
1	Get curriculum committee report	
2	Get sample curriculum materials	
3	Outline presentation	1,2
4	Write presentation	3
5	Type presentation	4
6	Make overhead transparencies	3
7	Give presentation	5,6

Figure 2-10.
Activities list with immediate predecessors filled in.

REVIEW

Terminology

1. PERT NETWORK: a pictorial representation of the interrelationships among the activities in a project.
2. EVENT: the beginning or ending point of an activity, represented in a network by a lettered circle.

Exercises

1. Using the activities list which you developed for the elementary school budget or the year-end staff evaluation project in Exercises 1 and 2, page 10, assign identification numbers to activities and construct a complete PERT network for the project.
2. Two activities lists for two different projects are given below. Draw the networks for each of these projects and add to the activities lists all correct immediate predecessors for each activity. Note that it is often helpful to construct a PERT network starting with the last activity and working backward to the first activity(ies).

(a) Activities list for project: Prepare a budget request for an elementary school.

ID	Activity Description
1.	Obtain class enrollment data.
2.	Obtain growth factor from past records.
3.	Predict next year's class enrollment.
4.	Determine number of classes.
5.	Determine size of teaching staff needed.
6.	Determine number of specialists needed.
7.	Obtain teacher supply requests.
8.	Determine supply needs.
9.	Determine transportation needs.
10.	Prepare budget request.

(b) Activities list for project: Polling the district teaching staff on their attitudes toward differentiated staffing.

ID	Activity Description	ID	Activity Description
1.	Prepare questionnaire.	7.	Distribute questionnaire.
2.	Prepare memo to staff regarding poll.	8.	Collect completed questionnaire.
3.	Type memo.	9.	Tally results.
4.	Distribute memo.	10.	Write report on poll.
5.	Type questionnaire.	11.	Deliver report to superintendent.
6.	Duplicate questionnaire.	12.	Inform teaching staff of results.

3. Select a reasonably limited project with which you are familiar and con-construct a PERT network for it.

The Dummy Activity

At this point you should have a good feeling for the basics of a PERT network—activities (represented by solid arrows) and events represented by lettered circles) accounting for the major stages of a project. One additional item is highly useful in laying out a diagrammatic network—the *dummy activity* (represented by a dotted arrow) which is an artificial activity used to prevent inconsistencies or confusions in networks. The dummy activity has three special uses, which we will consider one at a time.

First, wh[illegible]project has multiple beginning or ending act[illegible] a basic network would show such activiti[illegible]inning or ending independently both o[illegible] other and of any start or finish of the pr[illegible]s a whole. For example, Figure 2-9 shows [illegible]oard presentation project beginning with two independent activities (getting report and getting materials) and it unrealistically suggests that each starts spontaneously out of nowhere, without relation to the project as a whole. A look at Figure 2-11 shows the utility of inserting dummy activities in such cases—with the dummy activities in place, the network suggests clearly that once the project starts (event 0), the two beginning activities commence. In addition, were the project to end in two independent activities, they could realistically be connected by dummy activities to one terminating event. Hence, clarity and conformity to real-world decision making can be added by the dummy activity.

So that dummy activities can be individually identified, they are given identification numbers just as all other activities are. Dummy activities can be numbered as convenient—either in order as they appear in the network or in sequential order after the last activity number in the network, as in Figure 2-11.

The second use of the dummy activity is in instances where two separate activities have the same immediate predecessor and both im-

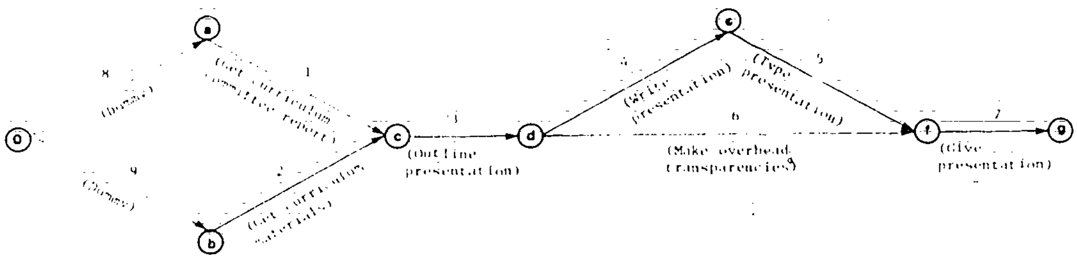


Figure 2-11.
Network for Board Presentation with Single Starting Event.

mediately precede the same activity. In such instances, the two activities share the same beginning event and ending event, and a simple PERT network would not be able to show the distinction between the two. A simple network could either show only one of the two activities (placing its identification number above the arrow between the two shared events) or it could show the two as indistinguishable (placing both identification numbers above the one arrow). The problem with the first alternative is obviously that of omission; the second alternative incorporates confusion.

In such cases, dummy activities can be used to distinguish the two activities clearly while showing them both to begin and end with a common event. This can be done by first drawing two separate arrows from the common beginning event and giving them separate ending events (one being the common event they both end in and the other being wholly artificial). These distinct ending events can then be diagrammatically shown to coincide by inserting a dotted arrow (dummy) between the artificial ending event and the actual ending event both activities precede.

To illustrate, suppose that the board presentation project required that while you are writing your presentation and having it typed and while your overhead transparencies are being prepared, you also must prepare a brief written summary of your presentation for the individual board members. Here the two activities of preparing overheads and preparing a summary both begin with the same event (the termination of outlining the presentation) and end with the same event (the start of making the presentation). Figure 2-12 shows how a dummy activity can be incorporated into the network so as to diagram the situation clearly. Beginning at event *d* (finish of outline), both activities are

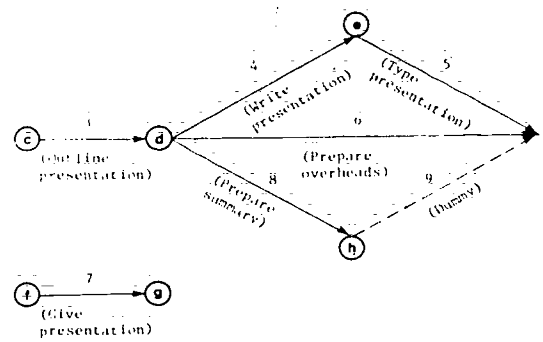


Figure 2-12.
Using a Dummy Activity to Distinguish Two Activities.

shown to proceed separately to event *f* (start of making presentation), one preceding directly and the other proceeding to its ending event *h*, which is then connected by a dummy activity to event *f*.

Dummy activities have a third use in networks—indicating a preferred sequence of activities when there is no other need for the specific sequence. Consider the board presentation project again and suppose that, in addition to the new activity of preparing a summary for the board, you must also plan to distribute this summary to the board. Furthermore, suppose that you want the activity of writing the presentation to precede the distribution of the summary, but there is no activity linking the two activities so as to establish the order of events. To make the desired precedence clear in the network, a dummy activity can be inserted, as illustrated in Figure 2-13, where the dummy activity connects events *e* (end of writing presentation) and *h* (beginning of distribution of summary).

When a network has been completed with dummy activities, the dummy activities are added to the complete activities list for the project. The immediate predecessors for each

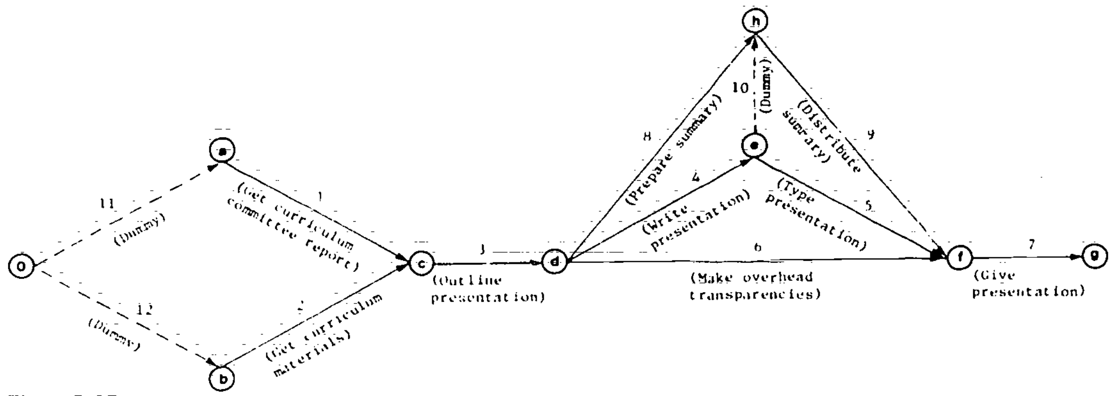


Figure 2-13.
Network for Board Presentation with Dummy Activity Indicating Time Precedence.

dummy activity are listed appropriately on the activities list.

In summary, the convention of the dummy activity is a valuable aid in drawing up complete and coherent PERT networks, whenever there is a need to:

1. Provide single starting or termination events for projects with multiple starts or finishes.
2. Represent two or more separate activities which have identical starting and ending events.
3. Indicate a desired precedence of activities where there is no intervening activity necessitating the sequence.

REVIEW

Terminology

1. DUMMY ACTIVITY: an artificial activity represented by dotted lines to prevent inconsistencies or confusions in PERT networks.

Exercises

1. Redraw the PERT network for the budget request project given in Exercise 2a on p. 14, using dummy activities appropriately.
2. Redraw the network for the polling project, Exercise 2b on page 14, using dummy activities.

III. DETERMINING COMPLETION TIME AND AVAILABLE SLACK

Once the network has been drawn and the predecessors clearly established, the next crucial issues concern time—how long each activity will take to complete and how long the entire project will take to complete. The next step is, then, to establish the duration of each individual activity. Once these durations are known, we can calculate the total amount of time it will take to complete the project. In addition, we can calculate how much delay can be absorbed by each activity without delaying the project. In the next section, we will explain how these calculations can be further used to determine the "critical path" in the project—the path or sequence of activities which is critical insofar as it has the least or no room for delay.

Activity Durations

The length of time (Duration) we must determine for each activity is the total length of time involved from beginning to completion, which often involves more time in preparation of the task than in actually doing it. For example, though it may take only one hour to make copies of a report, the entire activity might well take one or more days to complete if time must be allowed for getting the material to the copier, waiting for copying, and getting the material back from the copier.

For many projects, the length of time each activity will involve is well known from past experience. However, most educational administrators find themselves confronted regularly with projects involving activities for which durations must be estimated. In such cases, one rule has become cardinal—when collecting data to estimate activity durations, go to those most directly connected with executing the activity. Whenever this is not done or cannot be done, there is a high probability that the figures derived from the analysis of the PERT network will be worthless.

Once the activity durations have been determined, it is often convenient to enter them directly on the activities list. Durations can be entered in terms of hours, days, weeks, etc., depending on the nature of the project. Of course, the same Duration measure should be used for all activities within a project. As an example, the Duration for each activity in the board presentation project is shown on the activities list in Figure 2-14. In this case, all durations are shown in terms of days. Notice that the dummy activities that were established in the board presentation network in Figure 2-11 are entered on this activity list and are assigned durations of 0.0 days. Since dummy activities involve no time, 0.0 is the correct Duration in all circumstances.

When activity durations are established, they are entered on the PERT network under the arrow representing the appropriate activity. As an example, if an activity number 1 were "to get curriculum committee report" and its Duration were established as 1/2 day, it would be shown on the network like this:

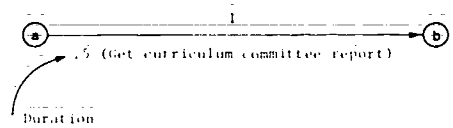


Figure 2-15.
Example of Placement of Activity Duration on Network.

The entire network for the board presentation project, with durations entered, is illustrated in Figure 2-16.

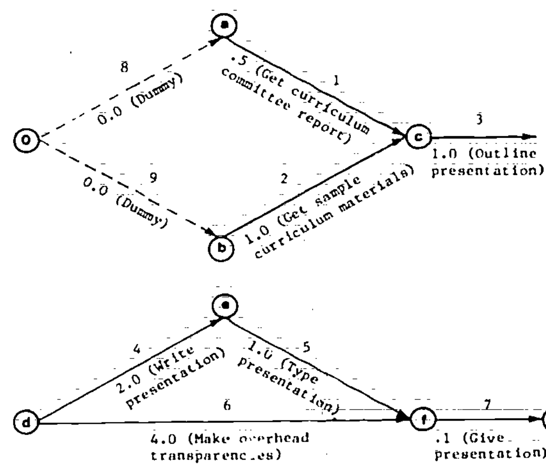


Figure 2-16.
Board Presentation Project with Duration Times.

ID	*Activity Description*	*Immediate Predecessors*	*Durations (days)*
1	Get curriculum committee report	8	0.5
2	Get sample curriculum material	9	1.0
3	Outline presentation	1,2	1.0
4	Write presentation	3	2.0
	Type presentation	4	1.0
6	Make overhead transparencies	3	4.0
7	Give presentation	5,6	.1
8	Dummy		0.0
9	Dummy		0.0

Figure 2-14.
Activities list with durations for board presentation project.

Project Completion Time

Earliest Start and Finish Times. With all activity durations entered, the network can be analyzed to determine the project Completion Time (CT). This analysis is carried out simply by determining the earliest start and finish times for each activity in sequence through the final activity.

The Earliest Start Time is calculated for each activity beginning with the initial starting time of zero for the project. Thus the Earliest Start Time for all initial activities is 0. The Earliest Finish Time is the Earliest Start Time plus the duration of the activity, or:

Earliest Finish = Earliest Start + Duration

For example, if an activity number 1, with an Earliest Start Time of 0 (zero), had a Duration of 1.5 (days), then its Earliest Finish Time would be 1.5 (days).

If the initial activity is always given the Earliest Start Time of 0, how do you suppose the Earliest Start Times of the subsequent activities are calculated? The Earliest Start Time of each subsequent activity is, of course, simply the latest Earliest Finish Time for all predecessors for that activity:

Earliest Start Time
= Latest Earliest Finish Time for Predecessors

For example, if activity number 16 is immediately preceded by activity 12 with an Earliest Finish Time of 4.5 (days) and by activity 15 with an Earliest Finish Time of 5.5 (days), then the earliest point at which activity 16 could begin would be 5.5 days—the latest Earliest Finish Time of all predecessors.

From the above discussion, you can see that the Earliest Start Times and Earliest Finish Times for each activity reflect the total amount of elapsed time from the start of the project's initial activity to the start or finish of the activity in question.

Each Earliest Start Time and Earliest Finish Time is entered on the PERT network for easy reference while the network is being further analyzed. These times are conventionally entered in brackets on top of the appropriate activity arrow, immediately after the activity's identification number. Earliest Start is entered first and Earliest Finish entered second, separated by a semicolon. For example, if an initial activity was "get curriculum committee report" with an activity Duration of .5 (days), it would have an Earliest Start Time of 0 and an Earliest Finish Time of .5, which would be entered as in Figure 2-17.

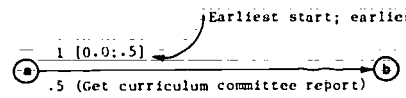


Figure 2-17.
Example of Placement of Earliest Start and Earliest Finish Times.

Figure 2-18 shows part of the network for the board presentation project with the Earliest Start and Earliest Finish times for the two initial dummy activities and activities 1, 2, and 3 correctly entered. Let's follow this example through, step by step.

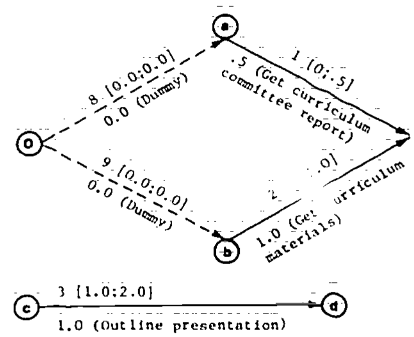


Figure 2-18.
Earliest Start and Finish Times for Initial Dummy Activities and First Three Acitivities on the Board Presentation Project Network.

Since we assume the starting time for the project is day 0, and since the dummy activities have duration times of 0, the Earliest Start and Finish times for both dummy activities in Figure 2-18 are 0. The Earliest Start times for activities 1 and 2 are then also 0. Since activity 1 has a duration of .5 days, then its Earliest Finish is .5 days after the project has begun. Since activity 2 has a Duration of 1.0 days, its Earliest Finish is 1.0.

Activity 3, however, must begin after both activity 2 has a Duration of 1.0 days, its Earliest

That is, the curriculum committee report (activity 1) and sample curriculum materials (activity 2) must be obtained before the outlining of the presentation (activity 3) can begin. We can be that sure activity 1 and activity 2 have been completed by day 1.0 (the Earliest Finish of activity 2). Therefore, the Earliest Start for activity 3 is day 1.0. Since activity 3 has a Duration of 1.0 days, then its Earliest Finish is day 2.0.

After calculating these quantities for each activity, one proceeds to derive the Completion Time (CT), or the Earliest Finish Time for the whole project, which is very simple. The Earliest Finish Time possible for a project is, naturally, the latest Earliest Finish of all activities and is therefore the same as the Earliest Finish of the project's terminating activity.

The entire network for the board presentation is shown in Figure 2-19. Earliest Start and Finish Times have been calculated for all activities. The total time for the project is CT = 6.1 days (Earliest Finish of activity 7).

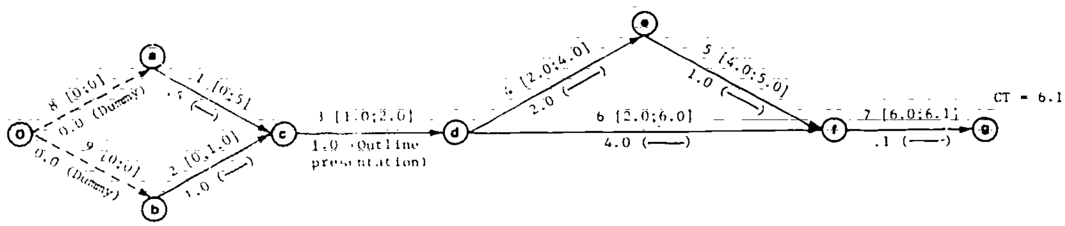


Figure 2-19.
Network for Board Presentation Project Showing Earliest Start and Finish Times and Completion Times.

REVIEW

Terminology

1. ACTIVITY DURATION: the total time it takes to complete an activity.
2. EARLIEST START TIME: the earliest time in a project at which a given activity can begin; for initial activities, always 0 (zero) and for other activities, the latest Earliest Finish for all predecessors to that activity.
3. EARLIEST FINISH TIME: for any activity in a project, the Earliest Start Time plus the Duration of the activity.
4. COMPLETION TIME (CT): the latest Earliest Finish Time of all the activities in a project.

Exercises

1. Beginning below are the complete activities lists (including activity Durations) for the budget request and the polling projects for which you drew networks in Exercises 1 and 2 on page 16. Using the networks that you previously drew, insert the Durations and the Earliest Start and Finish Times for each activity, and determine CT, the earliest Completion Time for each project.
   a. Activities list for project: Preparing a budget request for an elementary school.

ID	Activity Description	Immediate Predecessors	Durations (days)
1	Obtain class enrollment data	11	2.0
2	Obtain growth factor from past records	12	.1
3	Predict next year's class enrollment	1,2	1.0
4	Determine number of classes	3	.1
5	Determine size of teaching staff needed	4	.1
6	Determine number of specialists needed	5	.1
7	Obtain teacher supply requests		5.0
8	Determine supply needs	7	1.0
9	Determine transportation needs	13	3.0
10	Prepare budget requests	6,8,9	5.0
11	Dummy		0.0
12	Dummy		0.0
13	Dummy	12	0.0

b. Activities list for project: Polling the district teaching staff on their attitudes toward differentiated staffing.

ID	Activity Description	Immediate Predecessors	Durations (weeks)
1	Prepare questionnaire	14	1.0
2	Prepare memo to staff regarding poll	1	.2
3	Type memo	2	.1
4	Distribute memo	3	.6
5	Type questionnaire	1	.2
6	Duplicate questionnaire	5	.2
7	Distribute questionnaire	4,6	.6
8	Collect completed questionnaire	7	1.0
9	Tally results	8	.4
10	Write report on poll	9	1.0
11	Deliver report to superintendent	10	.1
12	Inform teaching staff of results	10	.6
13	Dummy		0.0
14	Dummy		0.0
15	Dummy	12	0.0
16	Dummy	11	0.0

Slack Time

The question now arises: How much can each activity in the project be delayed without delaying the whole project? In order to answer this question the latest possible time that each activity may be started and finished must be calculated. Once these latest times are established, we can compare them to the Earliest Start and Finish Times to see how much each activity may be delayed without delaying the on-time completion of the project.

Latest Start and Finish Times. The Latest Start and Finish times are calculated from right to left or backward along the network, beginning at the Completion Time, T, which cannot be delayed.

The Latest Finish for the activities which immediately precede the ending event of the project is, of course, the Completion Time, T. The Latest Start for any activity is:

Latest Start = Latest Finish
- Duration of Activity

For all activities other than those that precede the ending event, the Latest Finish is the Earliest Latest Start of all immediate successor activities.

As shown in Figure 2-20 below, this information is entered in brackets next to the activity duration figure. Notice that at this point in the network analysis we have removed the parenthetical activity description from below the activity line to make room for the Latest Start and Finish Times. Henceforth, the activity descriptions will be omitted from our networks; to identify each activity we can rely on the activity identification number always given above each activity arrow.

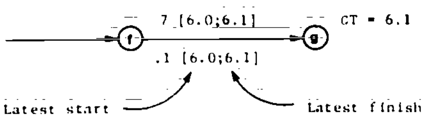


Figure 2-20.
Placement of Latest Start and Finish Times.

Figure 2-21a illustrates the network for activities 5, 6, and 7 of the board presentation project, showing Latest Start correctly.

The calculations for these times were made in the following way. First, the Latest Finish for activity 7 (giving the presentation) is equal to the finishing time of the project, T, or 6.1 days. The Latest Start of activity 7 is the Latest Finish of that activity minus the duration of the activity, or 6.0. A similar procedure is employed for each activity in the network, working from right to left.

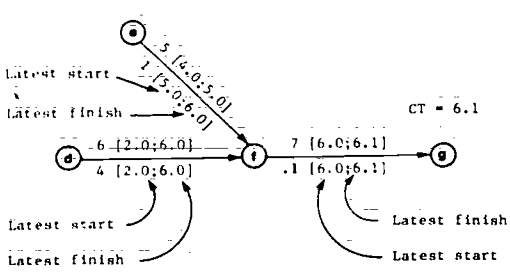


Figure 2-21a.
Latest Start and Finish Times for Activities 5, 6, and 7 of Board Presentation Project Network.

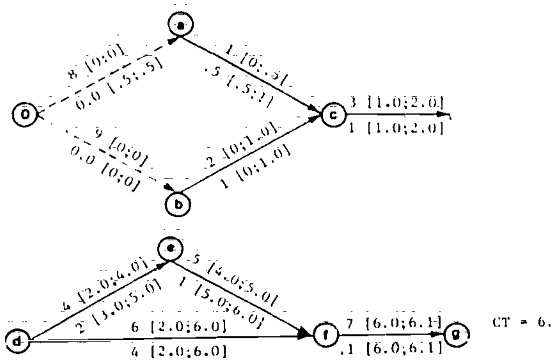


Figure 2-21b.
Board Presentation Project Network with Latest Start and Latest Finish Times.

Figure 2-21b below illustrates the placement of Latest Start and Latest Finish Times on the entire board presentation project network.

Let's consider the calculations for activity 3 (outlining the presentation) in detail. The two successors to this are activities 4 and 6. Their Latest Start figures are 3.0 and 2.0 respectively. Thus, since the Latest Finish of activity 3 is the minimum of the Latest Start Times for the immediate successors of activity 3, the Latest Finish of activity 3 is 2.0 and the Latest Start is the Latest Finish minus the duration of the activity, or 2.0 - 1.0 = 1.0.

Slack. From the Earliest and Latest Time figures, we can now make one final calculation which will establish how much delay can be absorbed up to any point in the project. This is the calculation of the Slack, or S, for each activity. Slack for an activity is defined as the amount of time which an activity may be delayed in either start or finish without delaying the completion of the project.

The Slack for an activity is the difference between the Latest Start and the Earliest Start, or the difference between the Earliest and Latest Finish Times for that activity.

Slack = Latest Start - Earliest Start
or
= Latest Finish - Earliest Finish

Slack is calculated for each activity beginning with the last one on the network and moving backward to the initial one(s) of the project.

Let's use the board presentation project network as an example and calculate the Slack for

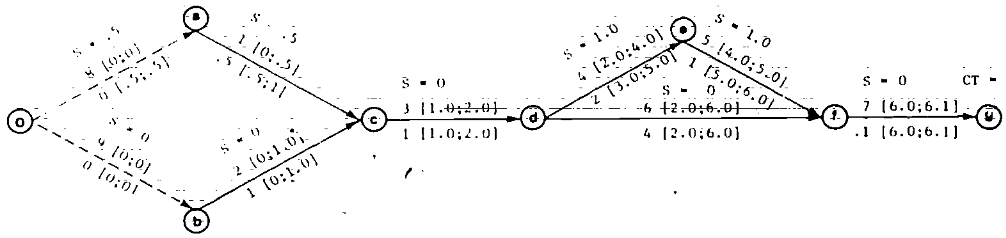


Figure 2-22.
Network for Board Presentation Project with Slack(s) Figures.

each activity, one by one. From the complete network shown in Figure 2-22 above, you can see that the Latest Start and the Earliest Start of activity 7 (giving the presentation) is 0 (zero); for this activity, the slack is therefore 0 and is entered as S = 0 above the appropriate arrow.

S for activity 6 is also zero, since the Earliest Start is again equal to the Latest Start. However, for activity 5 (typing the presentation), S = 5.0 – 4.0 = 1.0. Continue on through the remainder of the network and verify the Slack figures shown for each activity. As you can see, some activities have slack and others do not.

Once the S for each activity has been established, it is important to look at the meaning of the slack time for each individual activity and for sequences of activities. For example, while activity 4 (between *d* and *e*) has an S of 1 day and activity 5 (between *e* and *f*) has an S of 1 day, these actually represent the same day. If activity 4 is delayed by a day, there will be no slack for activity 5.

REVIEW

Terminology

1. LATEST FINISH TIME: for activities immediately preceding the terminating event of a project, Latest Finish Time is always CT, the Completion Time; for all other activities, Latest Finish is earliest Latest Start of all immediate successors.
2. LATEST START TIME: the Latest Finish Time minus the Duration for that activity.
3. SLACK: total delay time that can be absorbed by activities without delaying the completion of the project as a whole.

Exercises

1. Using the networks for the budget request project and the polling project for which you calculated Earliest Start and Finish Times in Exercises 1a and 1b, pages 19 and 20, calculate and enter the Latest Start and Latest Finish Times for each activity.
2. Using the networks from Exercise 1 above, calculate and enter the Slack for each activity.
3. Assuming that the budget request project from Exercises 1 and 2 above must be completed by June 30 next year, what is the latest date the project can start in order for it to be completed on time?
4. Select a relatively limited project with which you are familiar, such as preparing for and holding a board meeting or starting a new interscholastic athletic activity. Carry out a PERT analysis of the project, including the complete activities list, the network, and all time determinations.

IV. DETERMINE THE CRITICAL PATH

Now we have all the information we need to establish one of the most useful features of a PERT return: The critical path through the project. As we stated earlier, the critical path is a particular sequence of activities in the PERT network which determines the minimum amount of time it will take to complete the project. It is called the critical path because the completion of each activity within the estimated duration time is necessary in order to complete the project within the total estimated time (CT). If one traces any other sequence of activities or paths through the network, he or she will find that each one will take the same or less time to complete than the critical path. That is, there may be some activities in other paths which can be delayed in their starting or finishing without affecting the projected finishing time. The amount of this delay is something that we have just calculated—the total slack for each activity. Thus, it should be evident that the critical path has the least possible slack.

To determine the critical path for the board presentation example, look again at Figure 2-22. If we work backward through the network, we see that activity 7 has no slack. Of the two activities preceding activity 7, only activity 6 has no slack. Only activity 3 precedes activity 6. And of the two activities preceding activity 3, only activity 2 has no slack. The critical path is thus determined. It consists of the activities in the order 2 - 3 - 6 - 7.

The critical path is denoted by such a list of the activities on it. It is represented in the PERT network by double arrows. Figure 2-23 gives the complete network, including the critical path, which has resulted from a PERT analysis of the board presentation.

In the case we have been examining, there was no fixed time allotted for the project, but this is not always true. Many projects are allotted a given time period or required to be finished by a certain date. In this case, the critical path may not have zero slack. Suppose that in our board presentation example you decided before beginning the project to start two weeks before the meeting at which you were to give the presentation. What you would be doing, in effect, is to assign a value to CT of 10 days. Using this figure, the amount of slack in the network is increased for all activities, as can be seen

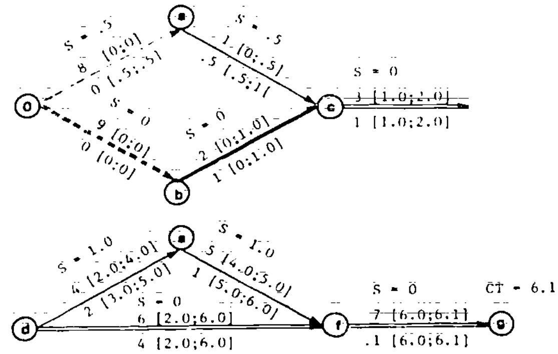


Figure 2-23.
Complete PERT Network for Board Presentation Problem.

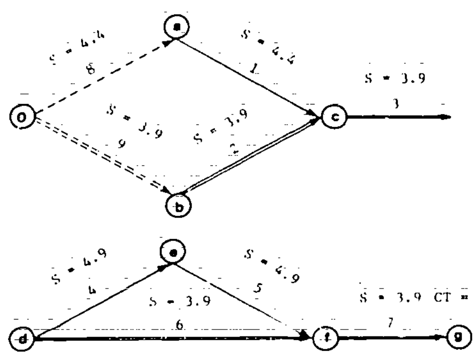


Figure 2-24.
PERT Network for Board Presentation Problem—Ten Days Allotted.

in Figure 2-24. *In this case, the critical path is the path through the network with the least amount of slack.* However, this obviously turns out to be the same path as before, 2 - 3 - 6 - 7.

When the time allotted for a project is less than the time required to complete the project, each activity on the critical path will have negative slack. This can be illustrated by assuming that we have scheduled the board presentation project to begin on Monday and the board meeting is on Friday of the same week. In effect, we have assigned a value of 5 days to CT. Using this figure, we can calculate the slack for each activity in the network; the results are given in Figure 2-25. The critical path is still that path through the network which has the least amount of slack[2]—again, it is 2 - 3 - 6 - 7.

The three cases that have been discussed above can be combined into one rule for finding

[2] Remember when working with negative numbers that -5 is a *smaller* number than -3.

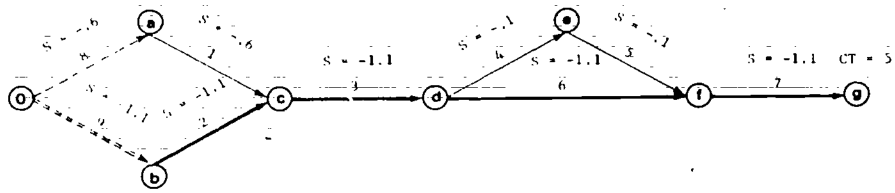


Figure 2-25.
Network for Board Presentation Problem—Five Days Allotted.

the critical path. The definition which covers all three cases is: *the critical path is that path through the network which has the least amount of slack.* Using this definition, it makes no difference whether the slack is positive or negative or zero. In all practical situations, however, if there is a negative slack in the critical path, either the project will automatically be completed after the deadline or some activities must be completed in less time than was originally estimated. In every case it will be up to the administrator to determine the most judicious course to follow given the particular circumstances, whether it is to start early, to cut down the time spent on some of the activities, or to allow the project to run over deadline.

For most projects, the administrator will obviously want to begin early if possible, in order to leave extra time throughout in case of unforeseen delays. This will amount to adding slack and, for purposes of the network, can simply be added to the slack shown for each activity. For example, if a given project has a Completion Time of 30 days and the administrator decides to start the project 40 days before the end date, then he would add 10 days to the slack figure for each activity.

REVIEW

Terminology

1. CRITICAL PATH: the path through a network which has the least amount of slack.

Exercises

1. (a) Use the networks for the budget request project and the polling project which you have been developing in the last several exercises and show on each the critical path.
   (b) Assuming that you have ten working days in which to carry out the budget request project, discuss what decisions you would make concerning the planning and implementation of the project.
2. Using the network you developed for your selected project in Exercise 4, page 22, determine the critical path and make any adjustments you deem necessary to carry out the project adequately.

SUMMARY OF PERT SYSTEM & THE CRITICAL PATH

Through the procedures outlined, it is possible to calculate the critical path for any project which consists of a set of discrete, interrelated activities for which times for completion are available. The critical path may be determined by using the following rules:

1. Construct an activity table which assigns

a numerical identifier to each activity, lists each activity's predecessors, and gives Durations for each activity.

2. Construct a PERT network of the activities and events.
3. In a left-to-right pass through the network let S = 0 and calculate:
   a. Earliest Start = S for beginning activities or
   b. Earliest Start = maximum Earliest Finish for all predecessors of the activity;
   c. Earliest Finish = Earliest Start + duration time;
   d. T = maximum Earliest Finish of all activities or the finish time assigned.
4. In a right-to-left or backward pass calculate:
   a. Latest Finish = T for all ending activities, or, Latest Finish = minimum Latest Start for all successors of the activity;
   b. Latest Start = Latest Finish - duration time of the activity;
   c. Total Slack = Latest Start - Earliest Start = Latest Finish - Earliest Finish.
5. Find the critical path by determining which sequence of activities or path through the network has the least amount of slack.

REVIEW

Exercise

Suppose you have just been given responsibility for conducting a school census. The date is now September 1. In order to plan and execute this census with maximum efficiency, you decide to use PERT. The table below gives the list of activities that you have decided will be in the project. Do the following tasks:

(a) Construct a PERT network of the project with all relevant data included and identify the critical path.

(b) Determine the starting date so that the report will be ready to present to the board by May 1. Assume that each week has five working days.

Activity List for Census Project

ID	*Activity Description*	*Duration (weeks)*	*ID*	*Activity Description*	*Duration (weeks)*
1	Determine information needed	2.0	9	Determine number of workers	.4
2	Design forms	3.0	10	Print forms	6.0
3	Determine census grids	2.0	11	Employ workers	2.0
4	Arrange computer processing	.2	12	Meetings with workers	1.0
5	Arrange publicity	.4	13	Take census	4.0
6	Order forms	.2	14	Keypunch data	3.0
7	Write computer programs	4.0	15	Computer analyses	.4
8	Publicity	4.0	16	Write final report	4.0

PERT ANALYSIS USING THE COMPUTER

PERT WITH PROGRAM GCPATH

The examples that have been discussed on the preceding pages have all been fairly simple in nature and thus were easily solved by hand. Many of the problems to which PERT is applicable, however, are not as easily solved. Often there are a large number of activities and complex relationships among the activities. PERT probably finds its greatest utility in those projects which encompass a great number of activities with numerous interrelations, but large projects like this are difficult to handle by the manual procedures we have outlined. Fortunately, a number of computer programs exist which do an efficient job of solving large PERT network problems. In this section we will introduce you to one of them.

The program we will use is named GCPATH and was written in the BASIC program language. The function of the program is to provide the PERT network analysis by taking in such initial information about the activities as Durations and predecessors and giving back the numerical figures for each. It will also identify which activities are on the critical path.

THE STEPS FOR USING GCPATH

Complete Activities List

Before beginning to use GCPATH, you will need a complete activities list, including all activities and dummy activities, predecessors, and Durations for all activities. To illustrate how GCPATH works, let's use again the board presentation project. The complete activities list for this project is reproduced in Figure 2-26 below.

Data Statements

While there are numerous different ways in general to make information or data available to a computer program for execution, GCPATH receives data by means of DATA statements which the user inputs into the computer. The form of this statement is:

statement number DATA activity ID,
Duration, predecessor, predecessor,..., -1

One DATA statement must be constructed for each activity in the project. Since there are seven activities in the board presentation project, we will need to construct seven DATA statements. Each must have a statement number; the first should be 2000.

Let us construct the first three DATA statements in detail:

```
2000 DATA 1,.5,-1
2001 DATA 2,1,-1
2002 DATA 3,1,1,2,-1
```

The first statement, numbered 2000, represents activity 1, which has a Duration of .5 days and has no predecessors. The -1 at the end of each line indicates that there are no more predecessors for this activity. Since each statement must be numbered with a consecutively higher number, the second is 2001. It indicates that activity 2 has a Duration of 1 day and has no predecessors. The DATA statement numbered 2002 indicates that activity 3 has a Duration of 1 day and, as predecessor, activities 1 and 2. Each following DATA statement is constructed in a similar manner until the list of activities for the project is exhausted. Figure

ID	*Activity Description*	*Predecessors*	*Durations (days)*
1	Get curriculum committee report	8	.5
2	Get sample curriculum material	9	1.0
3	Outline presentation	1,2	1.0
4	Write presentation	3	2.0
5	Type presentation	4	1.0
6	Make overhead transparencies	3	4.0
7	Give presentation	5,6	.1
8	Dummy		0.0
9	Dummy		0.0

Figure 2-26.
Complete activities list for board presentation project.

2-27 gives the complete list for the board presentation project.

```
2000  DATA 1,5,8,-1
2001  DATA 2,1,9,-1
2002  DATA 3,1,1,2,-1
2003  DATA 4,2,3,-1
2004  DATA 5,1,4,-1
2005  DATA 6,4,3,-1
2006  DATA 7,.1,5,6,-1
2007  DATA 8,0,-1
2008  DATA 9,0,-1
```

Figure 2-27.
Data statements for the board presentation project.

Getting Access to the Program

Once you have the list of DATA statements written down, you are ready to run the computer program GCPATH. The first thing you must do is to communicate to the computer that you want to run program GCPATH. To do this, you type the access statements appropriate for your computer.[3] For example, it may be:

```
GET-GCPATH
```

This statement entered into the computer will tell the computer that you want access to program GCPATH.

Entering Data and Running GCPATH

Next, you must enter your DATA statement for your project. In this case enter the DATA statements just as they are given in Figure 2-27. After you have typed each line, press the *return* key.[4] Once you have entered all the DATA statements, you are ready to run the program. To do this, type the run statement appropriate for your computer[5]—for example, RUN. The computer will reply by typing out the question:

```
HAVE YOU ENTERED YOUR DATA ALREADY? YES=1, NO=0
?
```

[3] If you have not operated a computer from a terminal before, consult your instructor for help.

[4] If you make an error in entering any DATA statement, simply hit the *return* key and retype the complete line correctly—e.g., 2000 DATG (error, hit *return*) 2000 DATA 1, .5, -1.

[5] Before beginning to use the terminal, be sure you know the access and run statements appropriate for your computer.

If the answer is yes, type "1" next to the question mark. If you should happen to type "0," the machine will print out the following lines:

```
ENTER THE PROJECT DESCRIPTION IN DATA STATEMENTS
BEGINNING WITH LINE 2000
FOR EACH JOB, GIVE THE FOLLOWING DATA --
  JOB NUMBER
  TIME REQUIRED TO COMPLETE
  PREDECESSOR JOBS (IF ANY)
  -1

JOBS MAY BE ENTERED IN ANY ORDER

AFTER ENTERING YOUR DATA STATEMENTS, RE-RUN THE PROGRAM
```

In essence, the computer is telling you how to enter your data. You would then enter your DATA statements and type RUN again. Using program GCPATH, the computer will now analyze your data. After it is finished it will type the information in figure 2-28.

```
ACTIV    DURATIONS    PREDECESSORS
-----    ---------    ------------
1        .5           8
2        1            9
3        1            1     2
4        2            3
5        1            4
6        4            3
7        .1           5     6
8        0
9        0

EARLIEST COMPLETION TIME FOR THE ENTIRE PROJECT = 6.1

         EARLIEST                LATEST
ACTIV    START    FINISH         START    FINISH       SLACK
-----    -----    ------         -----    ------       -----
1        0        .5             .5       1            .5
8        0        [illegible]    .5       .5           .5
[illegible] 0     1              0        1            0          *CP*
[illegible] 0     [illegible]    0        4            0          *CP*
3        1        [illegible]    1        2            0          *CP*
4        2        4              3        5            1
5        4        4              5        6            1
6        2        6              2        6            0          *CP*
7        6        6.1            6        6.1          0          *CP*

DONE
```

Figure 2-28.
Computer output for the board presentation problem.

The first information the program prints out (Part A) is a list of the activities in the project along with their duration time and list of the activities which precede them. This is printed out so that you may confirm that you have entered all the information on the activities of the project correctly. In Part B, the program prints out the earliest completion time of the project, or *T*—the minimum number of days it will take to complete the project. Finally, in Part C, it prints out for each activity the Earliest Start and Finish, the Latest Start and Finish, the Slack Time, and whether or not the activity is on the critical path. For instance, in Figure 2-28, activity 2 ("get curriculum materials samples") has an Earliest Start time of 0 days and Earliest Finish time of 1 day, a Latest Start time of 0 days and a Latest Finish time of 1 day, and a Slack of 0 days and is on the critical path. In contrast, activity 4 ("write the presentation") has an Earliest Start time of 2 days, an Earliest Finish time of 4 days, a Latest Start time of 3 days, a Latest Finish of 5 days, and a Slack time of 1 day; and it is not on the critical path. The critical path through the PERT network is thus identified by the activities that are labeled *CP* by the computer program.

COMPUTER PERT ANALYSIS: AN EXAMPLE

The chief advantage of computer solutions to PERT analyses is that they can provide time and slack figures and designate the critical path much more quickly and easily than hand solutions. In addition, the computer is capable of handling very complex projects with a large number of activities just as easily as it handled the small example above.

To illustrate the capability of GCPATH to handle a somewhat more complex project, let's look at one final example. Suppose that your state legislature passed a law during its last session requiring each school district in the state to establish educational goals for each grade level from first through twelfth within the next two years. Assume that you are in charge of this project for your school and that you come up with the activities list shown in Figure 2-29. From this list, you might construct a network similar to Figure 2-30.

This network would clearly illustrate predecessor activities and the position of dummy activities, and you could now complete the activities list by entering the latter, as in Figure 2-31.

If you had to rely solely on hand calculations at this point to derive Earliest and Latest Start and Finish Times, Slack, and the Critical Path, you would be in for a time-consuming set of operations. With GCPATH available, however, you could simply make your data statement list with activity number, Durations, and predecessors, and GCPATH would give you the full PERT analysis, including all activities on the critical path. Figure 2-32 shows the GCPATH printout results for the goal specification project.

ID	*Activity Description*	*Duration (weeks)*	*ID*	*Activity Description*	*Duration (weeks)*
1	Secure state guidelines	4	11	Generate preliminary set of goals	8
2	Construct committee guidelines	3	12	Distribute for teacher reaction	4
3	Appoint goal specification committees	4	13	Call in consultant	6
4	Contact state university education department regarding goal specification workshops	6	14	Compare with goals of other districts	2
5	Meet with workshop leaders	2	15	Revise goals	8
6	Schedule workshops	2	16	Submit to superintendent for approval	4
7	Distribute committee guidelines	4	17	Submit to school board for approval	8
8	Conduct workshops	9	18	Submit to state department of education for approval	10
9	Secure samples of goal specification from other districts	16	19	Publish, distribute to all teachers	14
10	Consult or review literature on goal specification	12	20	Inform community of outcomes	4

Figure 2-29.
Activities list for goal specification project.

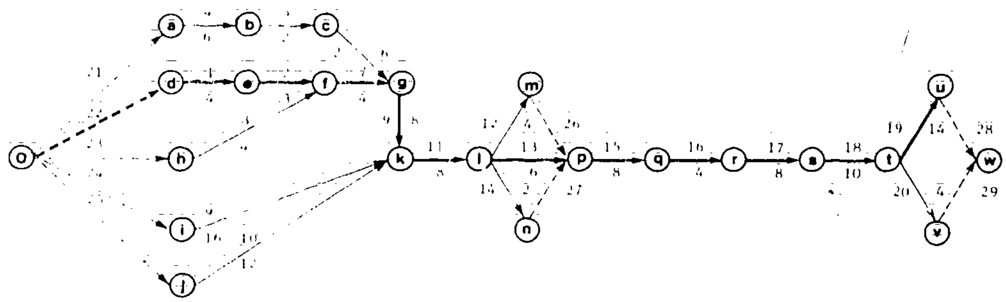


Figure 2-30.
Network for the Goal Specification Project.

ID	Activity Description	Predecessors	Duration (weeks)
1	Secure state guidelines	22	4
2	Construct committee guidelines	1	3
3	Appoint goal specification committees	23	4
4	Contact state university education department regarding goal specification workshops	21	6
5	Meet with workshop leaders	4	2
6	Schedule workshops	5	2
7	Distribute committee guidelines	2,3	4
8	Conduct workshops	6,7	9
9	Secure samples of goal specification from other districts	24	16
10	Consult or review literature on goal specification	25	12
11	Generate preliminary set of goals	8,9,10	8
12	Distribute for teacher reaction	11	4
13	Call in consultant	11	6
14	Compare with goals of other districts	11	2
15	Revise goals	26,13,27	8
16	Submit to superintendent for approval	15	4
17	Submit to school board for approval	16	8
18	Submit to state department of education for approval	17	10
19	Publish, distribute to all teachers	18	14
20	Inform community of outcomes	18	4
21	Dummy		0
22	Dummy		0
23	Dummy		0
24	Dummy		0
25	Dummy		0
26	Dummy	12	0
27	Dummy	14	0
28	Dummy	19	0
29	Dummy	20	0

Figure 2-31.
Durations.

```
2000 DATA 1,4,22,-1
2001 DATA 2,3,1,-1
2002 DATA 3,4,23,-1
2003 DATA 4,6,21,-1
2004 DATA 5,2,4,-1
2005 DATA 6,2,5,-1
2006 DATA 7,4,2,3,-1
2007 DATA 8,9,6,7,-1
2008 DATA 9,16,24,-1
2009 DATA 10,12,25,-1
2010 DATA 11,8,8,9,10,-1
2011 DATA 12,4,11,-1
2012 DATA 13,6,11,-1
2013 DATA 14,2,11,-1
2014 DATA 15,8,26,13,27,-1
2015 DATA 16,4,15,-1
2016 DATA 17,8,16,-1
2017 DATA 18,10,17,-1
2018 DATA 19,14,18,-1
2019 DATA 20,4,18,-1
2020 DATA 21,0,-1
2021 DATA 22,0,-1
2022 DATA 23,0,-1
2023 DATA 24,0,-1
2024 DATA 25,0,-1
2025 DATA 26,0,12,-1
2026 DATA 27,0,14,-1
2027 DATA 28,0,19,-1
2028 DATA 29,0,20,-1


RUN
GCPATH

HAVE YOU ENTERED YOUR DATA ALREADY? YES=1, NO=0
?1

ACTIV      DURATIONS      PREDECESSORS
-----      ---------      ------------
  1           4            22
  2           3            1
  3           4            23
  4           6            21
  5           2            4
  6           2            5
  7           4            2      3
  8           9            6      7
  9           16           24
  10          12           25
  11          8            8      9      10
  12          4            11
  13          6            11
  14          2            11
  15          8            26     13     27
  16          4            15
  17          8            16
  18          10           17
  19          14           18
  20          4            18
  21          0
  22          0
  23          0
  24          0
  25          0
  26          0            12
  27          0            14
  28          0            19
  29          0            20

EARLIEST COMPLETION TIME FOR THE ENTIRE PROJECT = 78
```

Figure 2-32.
GCPATH Analysis of the Goal Specification Project.

	EARLIEST		LATEST			
ACTIV	START	FINISH	START	FINISH	SLACK	
1	0	4	0	4	0	*CP*
22	0	0	0	0	0	*CP*
2	4	7	4	7	0	*CP*
3	0	4	3	7	3	
23	0	0	3	3	3	
4	0	6	1	7	1	
21	0	0	1	;	1	
5	6	8	7	9	1	
6	8	10	9	11	1	
7	7	11	7	11	0	*CP*
8	11	20	11	20	0	*CP*
9	0	16	4	20	4	
24	0	0	4	4	4	
10	0	12	8	20	8	
25	0	0	8	8	0	
11	20	28	20	28	0	*CP*
12	28	32	30	34	2	
13	28	34	28	34	0	*CP*
14	28	30	32	34	4	
15	34	42	34	42	0	*CP*
26	32	32	34	34	2	
27	30	30	34	34	4	
16	42	46	42	46	0	*CP*
17	46	54	46	54	0	*CP*
18	54	64	54	64	0	*CP*
19	64	78	64	78	0	*CP*
20	64	68	74	78	10	
28	78	78	78	78	0	*CP*
29	68	68	78	78	10	

DONE

Figure 2-32. Continued

REVIEW

Exercises

Use GCPATH to analyze one of the problems listed below that you have done previously by hand.

1. Preparing a budget request for an elementary school from page 14.
2. Polling the district teaching staff on their attitudes toward differentiated staffing from page 14.
3. Conducting a census for your school from page 25.

Your printout will depend upon the exact data you included in your initial analysis of the project.

USING PERT IN THE EDUCATIONAL SETTING

ORGANIZING, PLANNING, AND TIME MANAGEMENT

As you can appreciate at this point, PERT provides an invaluable approach to the organization of thinking and planning for any project which involves a well defined collection of interrelated activities. Some examples of typical administrative projects appropriate for PERT methods of analysis are building a new building, introducing a new instructional program into the curriculum, reorganizing existing programs, budget preparation, teacher-salary negotiation processes, planning affirmative action programs, planning school integration, and conducting needs surveys.

Using the network sequence, subprojects can be broken out of the network and designated as specific units of work to be delegated. Also, using the time figures from the analysis, activities can be systematically scheduled on the calendar.

Once a project is under way, the PERT analysis can be used as a valuable guide in managing the progress of the project in terms of time. Not only can the PERT-derived schedules be used, but the critical path provides check points to which the manager may want to pay particular attention.

PERT AS A COMMUNICATION TOOL

Once a PERT analysis is complete, it provides ample material for communicating effectively with subordinates, superiors, and any others who may require information about the project. The network alone provides a graphic picture of the interconnection between activities and among the subprojects, giving workers a clear understanding of how their tasks fit into the overall picture. The specific time figures, including slack and critical path factors as well as the resulting calendar schedule, clarify the time constraints for all concerned.

The administrator can summarize the PERT analysis information in any number of ways in order to communicate effectively about the project, depending on the situation.

AVAILABILITY OF COMPUTER PROGRAMS FOR PERT ANALYSES

As with computer programs for other operation research techniques, most administrators find that they can rely on the director of their computer center or the service agency that provides computer processing to obtain computer programs for PERT analyses and to run them with the data provided by the administrator. Such programs are usually readily available from the manufacturer of the computer you have access to, from university computer centers, or from users' groups.

There is a great number of different PERT analysis programs, but they are all basically the same. They differ primarily in special details—for example, the exact form of the input or the features of the output may be somewhat different. Since PERT analyses always require basically the same initial data and produce basically the same PERT results, administrators will normally have little difficulty using whatever PERT program may be available.

ADVANTAGES AND LIMITATIONS OF PERT

Now that you have some idea of what PERT is and what it can do, it is time to step back and take the larger view of PERT as a management tool. Obviously, it will not solve every problem of school administration. Yet it does have advantages (as well as disadvantages) when employed as a decision-making tool.

Let us consider the advantages first. Wiest and Levy provide an excellent summary of the advantages of PERT/CPM[6] for management in general:[7]

1. . . . [PERT/CPM techniques] are useful at several stages of project management—from the early planning stages, when various alternative programs or procedures are being considered; to the scheduling phase, when time and resource schedules are laid out; and finally in the operational phase, when used as a control device to measure actual versus planned progress.
2. They are straightforward in concept and easily explainable to the layman with no background in network theory. Data calculations, while tedious for large projects, are not difficult. Basic critical path data may be hand calculated with reasonable speed for projects with up to 500 or 600 activities. Computer programs are readily available for larger projects.
3. The network graph displays in a simple and direct way the complex interrelations

[6] CPM refers to the independently developed technique of Critical Path Method, the essentials of which are incorporated in such general PERT analysis as that presented in this text.

[7] Jerome D. Wiest and Ferdinand K. Levy, *A Management Guide to PERT/CPM* (Englewood Cliffs, N.J.: Prentice-Hall, 1969).

of activities which comprise a project. Managers of various subdivisions of the project may quickly perceive from the graph how their portion affects, and is affected by, other parts of the project.

4. Network calculations pinpoint attention to the relatively small subset of activities in a project which is critical to its completion. Managerial action is thus focused on exceptional problems, contributing to more reliable planning and more effective control.
5. CPM enables the manager to reasonably estimate total project costs for various completion dates. These various trade-off possibilities, along with other decision criteria, enable him to select an optimum or near-optimum schedule.
6. CPM and PERT are applicable to many types of projects—from aerospace development projects to large construction and maintenance jobs, from new product introduction programs to missile countdown procedures. Moreover, they may be applied at several levels within a given project, from a single department working on a subsystem to multi-plant operations within a large corporation.
7. As simulation tools, they enable the manager to project into the future the effects of planned or unanticipated changes and to take appropriate action when such projections indicate the need for it. Thus, for example, the manager can quickly study the effects of crash programs and can anticipate in advance potential resource bottlenecks that might result from shortening certain critical jobs.

In addition, PERT contains several other advantages for the educational administrator. First, and perhaps foremost, PERT allows the person responsible for the project to examine the entire project systematically and to identify the activities involved and the relationships among them; stated briefly, it is an analytical tool for planning and analyzing projects. It benefits the administrator concerned with obtaining grants from a number of funding institutions in that many such institutions now require proposals to be accompanied by a management plan similar to PERT. Finally, PERT offers a highly useful method of documenting projects for future use, either in repeating the project or in planning new but similar projects.

On the other side, PERT has several limitations arising primarily from the assumptions it makes about projects.

First, PERT assumes that a project can be subdivided into a set of predictable, independent activities. This has been challenged on two grounds. The first is that it may not be possible for all the activities which comprise a new project to be known before the project is started. This is not unreasonable when a group of individuals engage in a new project which they know little about; but for school administrators, it is unlikely that many completely unfamiliar projects would be initiated. The second challenge has to do with the identification of separate activities within the project. It maintains that in some cases it is not possible to separate the activities into independent units. For instance, the managerial function in a project is difficult to fit into the definition of an activity as having definite beginning and ending points within the project. This objection may be valid for several types of activities.

The second PERT assumption is that the relationships among activities can be completely represented by a noncyclical network in which each activity connects directly into its immediate successors. The assumption is not valid where the relationship between activities is conditional, i.e. the case where the order of activities is contingent upon the outcome of other activities. In a curriculum development project, for example, publication of the materials may hinge on the successful outcome of the field trials. If the trials are not successful, the materials must be revised—otherwise they go to publication. PERT has no provisions for this type of conditional precedence of project activities.

The third assumption of PERT is that activity durations may be estimated and are independent of each other. Since the calculation of the critical path depends on the time estimates, it is necessary that the estimates be accurate. But this is not always the case: when a project is initiated, the personnel involved may know little about the duration of the activities they must carry out and their estimates may be inaccurate. There may also be cases where the activity durations are not independent. When there is only a limited amount of manpower or money available, for instance, applying it to one activity may result in another activity being slowed down or stopped.

Finally, PERT has one characteristic which may prove a disadvantage in some situations. This is that it generally takes some time to carry out the data gathering and calculations necessary for a PERT analysis. If the project is fairly small and the time schedule tight, it may be more efficient to do without the analysis. However, this saving in time must always be weighed against the ultimate cost of poor scheduling and planning for the project.

FINAL EXERCISE

Choose a project which is or might be your responsibility and perform a PERT analysis, using GCPATH to aid in your analysis. You should generate the following:

1. An activities list;
2. A PERT network of the project;
3. A PERT analysis by the computer;
4. A verbal description of the critical path, making any advisable adjustments;
5. Using actual calender dates, schedule the activities for your project.

BIBLIOGRAPHY AND SUGGESTED REFERENCES

Cook, D. L. *Better Project Planning and Control through the Use of System Analysis and Management Techniques.* ERIC ED019 729. (Paper presented at Symposium on Operations Analysis of Education, sponsored by the National Center for Educational Statistics, Washington, D.C., November 20-22, 1967.)

Federal Electric Corporation. *A Programmed Introduction to PERT.* New York: John Wiley and Sons, Inc., 1963.

Harris, Roy D., Michael J. Maggard, and William G. Lesso. *Computer Models in Operations Research.* New York: Harper and Row, 1974. Suggested reading: Chapter 4, CREW, pp. 69-88.

Siemens, Nicolai, C. H. Marting, and Frank Greenwood. *Operations Research,* New York: The Free Press, 1973. Suggested reading: Chapter 4, Planning Long-Range Objectives, pp. 65-103.

Van Dusseldorp, Ralph A., Duane E. Richardson, and Walter J. Foley. *Educational Decision-Making through Operations Research.* Boston: Allyn and Bacon, Inc., 1971.

Wiest, Jerome D., and Ferdinand K. Levy. *A Management Guide to PERT/CPM.* Englewood Cliffs, N.J.: Prentice-Hall, 1969.

3 LINEAR PROGRAMMING

CHAPTER PREVIEW

This chapter introduces the technique of linear programming as a management tool for decision making. Linear programming is a problem-solving technique useful in problem situations where one quantity must be optimized (made as large or small as possible) while other quantities must remain within certain limits. An example of such a problem is determining the amount of instructional time to be spent in each of several different teaching modes so that the value of instruction to students is maximized while the limits of teachers and facilities available are observed.

The chapter contains a sample of a linear programming computer program and illustrates its use in several different situations. At key points, exercises are given to provide the user with experience in using a computer program and applying specific techniques for handling various linear programming problems.

The problems used as examples and exercises in this chapter are real-world problems which have been simplified somewhat in order to illustrate clearly the various techniques of linear programming.

The authors wish to thank the editors of Allyn and Bacon, Inc., for permission to reproduce and adapt portions of chapter 5, "Linear Programming," Ralph A. Van Dusseldorp, D. E. Richardson and W. J. Foley, *Educational Decision-Making through Operations Research*, Boston: Allyn and Bacon, 1971.

CHAPTER AIMS

Successful completion of this chapter with its exercises should provide the reader with a basic understanding of what constitutes a linear programming problem, how to form a mathematical model of the problem, and how to use a computer program to find the solution. Once these skills have been mastered, the reader should be able to deal successfully with real-world linear programming problems in all their complexity.

INTRODUCTION TO LINEAR PROGRAMMING

WHAT IS LINEAR PROGRAMMING?

In the course of their careers, educational administrators encounter many problems like the following:

How can we adjust the salary schedule so that beginning teachers' salaries are as high as possible and still remain within the budget?
Which teachers should we hire in order to maximize the quality of teaching in the district?
How can we have racially balanced schools while transporting students a minimum distance from their neighborhoods?
How much time should be devoted to individualized instruction in mathematics so that students are receiving the maximum possible benefit?

Although the above problems deal with widely varied topics, all of them are concerned with allocating or assigning some *resource*. In the first question we are interested in allocating *money* for a salary schedule. In the second, we are concerned with assigning *teachers* to positions. The third is concerned with the assignment of students to *schools*, while the fourth deals with the allocation of *class time*.

The questions posed above also have another common characteristic. They are all concerned with *making some quantity as large or as small as possible (maximizing or minimizing)*: in the first question, the maximum salary of a beginning teacher; in the second, the maximum quality of teaching; in the third, the minimum distance students must travel; and in the fourth, the maximum possible benefit in mathematics instruction.

Problems like those outlined above, which are concerned with allocating or assigning some resource in order to maximize or minimize a certain result, are called *problems in optimization*. Optimizing a quantity can refer to either maximizing or minimizing it.

Many optimization problems can be solved using the operations research technique of linear programming. In the past, linear programming has been used principally to solve problems in the military, industry, and business. In recent years, however, the flexibility and problem-solving power of linear programming have proved it to be a highly effective tool for educational decision making.

The term "programming" in "linear programming" does not specifically refer to computer programming. Rather, it refers to expressing or modeling a real problem situation in mathematical terms. The term "linear" refers to a particular kind of mathematical relationship between the variables or quantities in the problem. The reason this relationship is called linear is that graphs of linear mathematical formulas are always straight lines. Each of the formulas in the model of a particular situation must have this linear relationship of its variables before the technique of linear programming can be applied. A precise definition of a linear formula is given in the following example, and a more complete discussion of linear equations appears in the last section in this chapter.

EXAMPLE OF A LINEAR PROGRAMMING PROBLEM

To get a first look at how the technique of linear programming works, let's use a hypothetical problem borrowed from the business world. Suppose a television manufacturer produces two models of television sets—portable and console. The profit on each portable set produced is $20, while the profit on each console model is $25. The maximum capacity of the assembly lines is 60 portables and 40 consoles per day. The plant employs 150 assembly-line workers, all of who can work on either the portables or the consoles. It takes two person-days to produce a portable set and three person-days to produce a console. We assume that all television sets produced will be sold. The problem is to determine the number of consoles and portables to produce each day in order to maximize the total profit.

At the outset, this appears to be a problem in optimization because it involves certain *resources* to allocate (assembly lines and workers) and something to be *optimized* (the total profit).

The first step in solving the problem using linear programming is to express it in mathematical symbols. Let us choose T to stand for the total daily profit, the variable p to stand for the number of portable sets manufactured per day, and the variable c to stand for the number of console sets manufactured per day.

Since each portable set yields a profit of $20 and each console set a profit of $25, the total daily profit can be expressed in the following equation:

$$20p + 25c = T$$

Since our goal is to find the number of portables and consoles that should be produced each day so the maximum possible profit is made, we want to find values to assign to p and c that yield as large a T as possible.

There are some special terms used in linear programming to identify various parts of the problem. The *measure of effectiveness* is the quantity we want to either maximize or minimize—in this problem, the total daily profit. The mathematical equation expressing the measure of effectiveness is called the *object function*. In this case, the object function is:

$$20p + 25c = T$$

From the above equation for T, it initially appears that the manufacturer could maximize his profit by producing only console sets c. However, the number of consoles that can be produced is limited by plant capacity to 40 per day. This condition is symbolically expressed by[1]:

$$c \leqslant 40$$

In linear programming a restriction like this is called a constraint. In terms of the problem, in finding values for c and p which will maximize the plant's daily profit, we are limited or constrained by the fact that only 40 consoles can be build per day.

Another constraint is the restriction that no more than 60 portable sets can be built per day. This condition is expressed by:

$$p \leqslant 60$$

which indicates that p, the number of portable sets that can be produced, is less than or equal to 60.

A third constraint is that it takes two person-days to build a portable set and three person-days to build a console set. When p portables are built, this production requires $2p$ employees each day. If c consoles are built, $3c$ employees are needed each day. Only 150 employees are available to build television sets each day, so the third constraint expressed in mathematical terms is:

$$2p + 3c \leqslant 150$$

That is, the total number of employees used to build portable and console television sets must be no larger than 150.

At this point, we have expressed the elements of the television production problem mathematically, in the form of three constraints and one object function. The problem's conditions are mathematically stated. These mathematical statements—constraints and object function—comprise a mathematical *model* of the problem.

We are now in a position to examine the formulas in our model and to determine if they are linear. If they are, we can conclude that the problem is a linear programming problem; that is, it can be solved using linear programming techniques.

A formula (either an equation or an inequality) is *linear* if the variables are only added and/or subtracted from each other; that is, variables may *not* be multiplied or divided by each other or by themselves. Real numbers may be added, subtracted, multiplied, or divided at any point in the formula without affecting its linearity. A formula with one variable appearing only once is automatically linear. Study the following examples of formulas and the discussions of whether or not they are linear.

$x + y = 0$	Linear—because the variables are added together.
$3x - 4y \leqslant ½$	Linear—because the variables are subtracted from each other. Notice that the numbers can be multiplied

[1] The sign $\leqslant$ means "less than or equal to," so $c \leqslant 40$ is read "c is less than or equal to 40" or "the number of consoles produced must be less than or equal to 40." The sign $<$ is read "strictly less than." If we had written "$c < 40$," we would have excluded the possibility that c could exactly equal 40. The sign $\geqslant$ (the reverse of $\leqslant$) means "greater than or equal to," and $>$ of course means "strictly greater than."

	by the variables ($3 \cdot x$ and $4 \cdot y$).
$\frac{x - y}{2} \geqslant z + 4w + 1$	Linear—because the variables are added or subtracted from each other. Notice that numbers can be divided into variables ($x - y/2$) and can stand alone (+ 1).
$a \cdot b = 14$	Not linear—because the variables are multiplied by each other.
$3x^2 \leqslant x + 2$	Not linear—because the variables are multiplied by each other (x^2 means $x \cdot x$, so x is multiplied by itself).
$\frac{4x + y}{z} = x$	Not linear—because the variables are divided by each other ($4x + y/z$).

Note: There is no limit to the number of variables allowed in a linear formula.

Let's look at the formulas in our model to see if they are linear:

$$p \leqslant 60$$
$$c \leqslant 40$$
$$2p + 3c \leqslant 150$$
$$20p + 25c = T$$

All of them either have variables which are added to each other or contain only one variable. Therefore, all four of our formulas are linear formulas and this is indeed a linear programming problem. We have, therefore, accomplished these steps:

1. We have determined that the problem, one of resources to allocate and a quantity to optimize, is an optimization problem and therefore is a candidate for the technique of linear programming.
2. We have formulated the model by
   a. identifying the controllable variables and the measure of effectiveness;
   b. identifying the constraints;
   c. representing the constraints and measure of effectiveness mathematically.
3. We have examined the model and determined that the formulas are all linear.

In summary, then, we have the following information for this problem:[2]

Statement in Problem	*Mathematical Expression of Statement*	*Type of Statement*
The daily profit should be maximized.	Subject to the above constraints, find p and c to maximize . . . $20p + 25c = T$	Object function
No more than 60 portable sets can be built per day.	$p \leqslant 60$	Constraint
No more than 40 console sets can be built per day.	$c \leqslant 40$	Constraint
A total of 150 employees are available; it takes 2 person-days to build a portable and 3 person-days to build a console.	$2p + 3c \leqslant 150$	Constraint

[2] It is tacitly assumed that the number of portables and the number of consoles are greater than or equal to zero, i.e. $t_1 \geqslant 0$ and $t_2 \geqslant 0$. It would make no sense to talk about producing a *negative* number of television sets! It is unnecessary to include the expressions $t_1 \geqslant 0$ and $t_2 \geqslant 0$ as const[illegible] in the mathematical model. When we use a comp[illegible] program to solve linear programming problems, we will find the program automatically assumes that controllable variables cannot have negative values.

REVIEW

Terminology

1. OPTIMIZATION—the maximization or minimization of a given quantity.
2. MEASURE OF EFFECTIVENESS—the quantity to be maximized or minimized.
3. CONSTRAINT—a restriction that imposes limits on a problem's solution.
4. OBJECT FUNCTION—the mathematical equation expressing the measure of effectiveness.
5. LINEAR—characteristic of a formula in which variables are only added to and/or subtracted from each other, thus resulting in the graph of the formula being a straight line.
6. LINEAR PROGRAMMING—a problem-solving technique useful for real-life problems in optimization which can be expressed by a mathematical model involving linear formulas.

Exercises

1. Choose variables to represent the following quantities for a school:
   number of classes in session at any one time;
   number of teachers;
   number of teacher aides;
   average teacher's salary;
   average teacher aide's salary;
   number of children in the school.

   Now, practice translating English statements into equivalent mathematical statements by expressing the following conditions symbolically:
   a. The number of classes in session at any one time is no larger than 65.
   b. The number of children is at least 250.
   c. The number of teachers plus the number of teacher aides is less than or equal to 75.
   d. The average teacher's salary is $2,000 more than the average teacher aide's salary.
   e. The number of teachers is between 35 and 40.
2. Which of the following formulas are linear?
   a. $5x = y$
   b. $2/3a - 3b + c \geqslant 0$
   c. $y^2 + 5y + 6 = 1$
   d. $.5c - d = e - 3$
   e. $2mn - 6m + 5n \leqslant 0$

FINDING A SOLUTION TO THE TELEVISION PRODUCTION PROBLEM

Once the mathematical model for the problem has been formulated, the search for a solution involves finding the best solution from among many possible ones. The best solution to the television production problem is the one that maximizes the expression $20p + 25c = T$ and thus results in the greatest possible total profit, T.

What values of p and c will make T as large as possible? Remember that we are limited in choosing p and c by the following constraints:

$p \leqslant 60$ (number of portables)
$c \leqslant 40$ (number of consoles)
$2p + 3c \leqslant 150$ (number of person-days)

Let's try to solve this problem by trial and error, merely to illustrate how possible solutions to a problem differ from the best solution. (We will subsequently learn to use the computer program LPRG to solve linear programming problems.)

Suppose we choose, arbitrarily, to manufacture 60 portables and 40 consoles. That is, we choose values for p and c of 60 and 40 respectively ($p = 60$ and $c = 40$).

There are now two things we want to know: Have we found a *possible* solution to the problem and is this the *best* solution? We know we have found a possible solution if our choices for p and c satisfy all the conditions, or constraints, of the problem. We have found the *best* solution when T, the total daily profit, is as large as it can possibly be. There are therefore three things we want to check, now that we have assigned arbitrary values to p and c:

1. Does our choice of p and c satisfy the constraints?
2. If the constraints are satisfied, what is the resulting daily profit, T?
3. Is this profit as large as it can possible be?

Using the arbitrary values $p = 60$ and $c = 40$, we now answer these questions.

1. Are the constraints satisfied by $p = 60$ and $c = 40$?
   First constraint: $p \leqslant 60$.

   $$p = 60 \text{ and } 60 \leqslant 60;$$

   therefore, this constraint is satisfied.
   Second constraint: $c \leqslant 40$.

   $$c = 40 \text{ and } 40 \leqslant 40;$$

   therefore, this constraint is satisfied.
   Third constraint: $2p + 3c \leqslant 150$.
   Substituting $p = 60$ and $c = 40$, we have $(2 \cdot 60) + (3 \cdot 40) = 120 + 120 = 240$. 240 is not less than or equal to 150; 240 is greater than 150. Therefore, $p = 60$ and $c = 40$ is not a solution because it does not satisfy all the constraints.

Since we have learned we cannot use values for p and c as large as 60 and 40 respectively, let's try $p = 0$, $c = 40$ for a second possible solution.

1. Are the constraints satisfied by $p = 0$ and $c = 40$?
   First constraint: $p \leqslant 60$.
   $p = 0$ and $0 < 60$ so $0 \leqslant 60$; therefore, this constraint is satisfied.
   Second constraint: $c \leqslant 40$.
   $c = 40$ and $40 \leqslant 40$; therefore, this constraint is satisfied.
   Third constraint: $2p + 3c \leqslant 150$.
   Substituting $p = 0$ and $c = 40$, we have $(2 \cdot 0) + (3 \cdot 40)$ and $0 + 120 = 120 \leqslant 150$. Therefore, this constraint is also satisfied.
   Since all the constraints are satisfied, we can conclude that $p = 0$ and $c = 40$ is a solution to the problem.
2. What is the resulting daily profit, T? The object function is $20p + 25c = T$. Substituting $p = 0$ and $c = 40$, we find $20 \cdot 0 + 25 \cdot 40 = 0 + 1000 = 1000 = T$, so the daily profit is $1000 if 0 portables and 40 consoles are produced.
3. Is T as large as it can be?
   We don't know yet. It may be possible to choose values for p and c that result in a larger value for T. Let's try another set of values for p and c and see what happens to the value of T.

Again we select values for p and c. We know that $c = 40$ is as large as allowable according to the second constraint. For the next solution, let's keep $c = 40$ and make p as large as possible, hoping to maximize T. With a little forethought we can select a value for p which will be as large as possible and which together with $c = 40$ will satisfy all the constraints.

1. Our first concern is to pick values for c and p which satisfy all constraints. Since we already have chosen the value of c as 40, which satisfies the constraints, we turn to p. Let's look at the constraints to get a clue to what value to select for p.
   First constraint: $p \leqslant 60$.

Value for p (portables)	*Value for c (consoles)*	*Are all constraints satisfied?*	*Value for T*
0	40	yes	\$1000
10	20	yes	700
10	40	yes	1200
10	50	no*	—
15	35	yes	1175
15	40	yes	1300
20	35	yes	1275
30	35	no†	—
30	30	yes	1350
40	20	yes	1300
50	10	yes	1250
60	10	yes	1450
70	10	no‡	—

*$c = 50$ violates the constraint that $c \leqslant 40$.
†$p = 30$ and $c = 35$ violates the constraint that $2p + 3c \leqslant 150$ ($2\cdot30 + 3\cdot35 = 60 + 105 = 165$).
‡$p = 70$ violates the constraint that $p \leqslant 60$.

Table 3-1.
Sample values for p *and* c, *television production problem.*

We already know from our earlier attempt that p cannot be as large as 60 when $c = 40$. So we know that any smaller value that we chose for p will at least satisfy this constraint.

Second constraint: $c \leqslant 40$.

Since we are staying with $c = 40$, we know the second constraint is satisfied regardless of which value is chosen for p.

Third constraint: $2p + 3c \leqslant 150$.

Since we have already decided that $c = 40$, we can use this fact to figure out what restrictions the third constraint has on the value of p, which we have yet to determine. Using $c = 40$, the third constraint becomes

$$2p + 3\cdot40 \leqslant 150, \text{ or}$$
$$2p + 120 \leqslant 150.$$

From this we can see that $2p \leqslant 30$, which means $p \leqslant 15$. Therefore, in order to satisfy the third constraint, we must choose a value for p which is not larger than 15.

Since we want p to be as large as possible, we chose $p = 15$. So $p = 15$ and $c = 40$ is our new proposed solution. Because of the way values for p and c were chosen, we can be assured that $p = 15$ and $c = 40$ is a solution, since these values satisfy all three constraints.

2. What is the total daily profit, T?

$$T = 20p + 25c = (20\cdot15) + (25\cdot40) = 300 + 1000 = \$1300$$

3. Is T as large as it can be?
   T is larger than it was when we tried values of $p = 0$ and $c = 40$, but we still don't know if T is as large as it can possibly be. Suppose we have tested several more combinations for p and c and have summarized our findings as in Table 3-1.

From Table 3-1 we can see that there are many solutions to the television production problem—that is, it is possible to produce the following number of television sets while satisfying the given constraints:

Portables (p)	*Consoles (c)*
0	40
10	20
10	40
15	35
15	40
20	35
30	30
40	20
50	10
60	10

The combination of $p = 60$ portables and $c = 10$ consoles produces a profit of \$1450. If we could check every possible solution to this problem, we would find that the combination of $p = 60$ and $c = 10$ is the only solution which produces a daily profit of \$1450, and furthermore, that \$1450 is the maximum profit we could obtain. We call the values $p = 60$ and $c = 10$ the *optimal solution* to this problem; i.e., the values $p = 60$ and $c = 10$ satisfy the original constraints of the problem while at the same time optimizing (in this case, maximizing) the total daily profit.

This problem typifies many problems in linear programming in that it has many solutions but only one *optimal* solution.[3] We

[3] Some linear programming problems will have no solution at all, while others have an infinite number of optimal solutions. See pages 65-73 in this chapter.

could find many values for p and c that satisfy all the constraints in the television problem, but only one set of values will produce the largest possible profit. (Don't worry at this point about how we know \$1450 is the largest possible profit, or how we know that $p = 60$ and $c = 10$ is the optimal solution. We will later verify it through use of a computer program.)

To sum up, in solving the television production problem we first expressed the conditions of the problem as three constraints:

$$p \leqslant 60;$$
$$c \leqslant 40;$$
$$2p + 3c \leqslant 150.$$

We then found p and c such that the object function:

$$20p + 25c = T$$

was a maximum. The optimal solution was:

$$p = 60 \text{ portables}$$
$$c = 10 \text{ consoles}$$

The maximum value of T was \$1450.

The trial-and-error technique was used to illustrate the difference between the many possible solutions to a problem and the optimal solution. It is fairly easy to guess the optimal solution to simple linear programming problems like this one, but most practical linear programming problems contain far too many variables for the trial-and-error method to be useful or even possible. Our next step, then, is to learn how to solve linear programming problems using the computer program LPRG.

REVIEW

Exercises

1. What is the difference between a possible solution to a linear programming problem and the optimal solution? (Review text pages 39–42.)
2. Read the following problem, then answer the five questions below.

A school system wants to hire a total of 85 teacher aides for 85 elementary school classrooms. One aide will be permanently assigned to each classroom. Ten teachers have indicated that they definitely need full-time teacher aides. The rest (75) would prefer full-time aides if possible but would accept part-time aides. The school system, in an effort to improve community relations, is committed to hiring at least 15 qualified parents as part-time teacher aides. An additional consideration in hiring aides is that the school system wishes to qualify for additional federal funds by increasing staff. To meet the federal requirement, the school must employ at least 50 teacher aides as staff members. Full-time aides are counted as one staff member, part-time aides as one-half. The school system wishes to satisfy all the above conditions while at the same time spending the least amount of money. A part-time aide is paid \$7 a day, a full-time aide \$15 a day.

How many full-time and part-time teacher aides should be hired?

(a) What makes this problem a candidate for the linear programming technique of problem solving?

(b) Which quantities will be the two controllable variables in the model of the problem? Choose variable names for these quantities.

(c) What is the quantity we wish to optimize in this problem? Do we wish to maximize or minimize this quantity? Write this quantity mathematically as the object function.

(d) List the constraints on the controllable variables as concisely as as you can; then express them mathematically. (There are four constraints.)

(e) Set up a table similar to the one we constructed for the television problem and try to guess the optimal solution by varying the values for the controllable variables.

SOLVING LINEAR PROGRAMMING PROBLEMS BY COMPUTER

THE COMPUTER PROGRAM LPRG

The program we will use to solve linear programming problems is named LPRG. Since you are now familiar with the television problem, we will use it as a sample problem to illustrate how to use LPRG. In the last section of this chapter you will examine many examples of more complex problems relevant to the field of educational administration and study their solutions as given by LPRG.

How to Use LPRG

The following four steps are necessary to solve linear programming problems using LPRG.

1. *Completely formulate the mathematical model for the problem.* Write out in mathematical symbols all the constraints and the object function.

Recall that the mathematical model of the television problem consisted of the following object function and constraints, where p represents the number of portables and c the number of consoles:

$$\left.\begin{array}{r} p \leqslant 60 \\ c \leqslant 40 \\ 2p + 3c \leqslant 150 \end{array}\right\} \text{Constraints}$$
$$20p + 25c = \text{T} \;\} \text{ Object function}$$

We want to find values for p and c such that T, the total daily profit, is maximized.

Once all the constraints and the object function have been formulated, you will have to assign an artificial order to the controllable variables. It does not matter what the order is or what method is used to determine it. All that is important is that there is a first variable, a second variable, a third one, and so on.

In our example, the controllable variables are p (the run of portables) and c (the number of consoles). Let us agree (arbitrarily) that p is the first variable and c is the second.

Examine the first constraint. It says that p is less than or equal to 60, but it says nothing about c. When we enter information about the mathematical model of a problem into the computer using LPRG, we must make sure that *each* constraint says something about *each* controllable variable. Suppose we rewrite the first constraint as:

$$p + 0c \leqslant 60$$

We have not changed the meaning of the constraint at all. In effect, we have just added zero to the left-hand side, in the form of the variable c with a coefficient of zero.[4] Suppose, furthermore, that we write this constraint again, this time clearly indicating the coefficient of p is 1:

$$1p + 0c \leqslant 60$$

We must rewrite all the constraints and the object function of the mathematical model in this manner, clearly indicating what the coefficients of the variables are, and including *every* variable in *every* constraint. The purpose of rewriting the model is to facilitate entering correct data into the computer program LPRG.

Here is the revised (but equivalent) model for the television problem:

$$\left.\begin{array}{r} 1p + 0c \leqslant 60 \\ 0p + 1c \leqslant 40 \\ 2p + 3c \leqslant 150 \end{array}\right\} \text{Constraints}$$
$$20p + 25c = \text{T} \;\} \text{ Object function}$$

Notice that we have lined up the equations so all p's and c's are directly under each other in a column.

The constraints must also be in order. LPRG is organized to expect all the constraints involving $\leqslant$ first, the constraints involving = next,

[4] The *coefficient* of a variable is the number which multiplies it. The coefficient of $2p$ is 2; the coefficient of $3c$ is 3; the coefficient of p is 1 (p is one p); the coefficient of c is 1 (one c).

and the constraints involving ≥ last. If there is more than one constraint involving ≤, their relative order is unimportant and may be assigned arbitrarily. The same holds for two or more = or ≥ constraints.

In our example, all the constraints are of the ≤ variety, so their order is assigned arbitrarily. Since we have already distinguished $1p + 0c \leq 60$ as the first constraint, $0p + 1c \leq 40$ as the second, and $2p + 3c \leq 150$ as the third, we will retain this order.

The model is now in a form from which we can easily enter all the necessary information into LPRG.

2. *Run LPRG.* To run the computer program LPRG, first access the program by typing in the message GET-LPRG.[5]

When the computer responds by spacing up a line (or by typing READY or otherwise indicating that you are to proceed), type the command to start execution of the program: RUN.

The first line the computer types will identify the name of the program: LPRG.

Now you begin to tell the computer about your linear programming problem. You do this by responding to a series of questions which will be typed by the computer. The questions will be easy to answer when the problem is organized as the television set problem is organized in Figure 3-1. Let's begin:

```
IF MAXIMIZING THE OBJECT FUNCTION, TYPE '1';
IF MINIMIZING THE OBJECT FUNCTION, TYPE '-1'.
?1
```

Since we are maximizing the television set problem, we respond with a 1. The program continues:

```
NUMBER OF VARIABLES?
?2
```

This question refers to the controllable variables only. In the television problem we have two controllable variables, *p* and *c* (the number of portables and consoles).

[5] The message used to access LPRG may be different on the computer system you are using, and the command to begin execution may also be different, so be sure you know the correct commands before continuing. Notice that in this and other discussions involving interactions with the computer, all responses typed by the user will be underlined to distinguish them from computer-generated response. On actual runs, however, no underlining will appear.

```
NUMBER OF CONSTRAINTS?
?3
```

We have listed three constraints in Figure 3-1; therefore, the correct response is 3.

Next, you will be asked how many of the constraints are "less-than" (involving ≤), how many of them are "equality" (involving =), and how many of them are "greater-than" (involving ≥). You must account for all the constraints as one of these. If you do this incorrectly the program will tell you the data you entered is inconsistent and you will have to enter it again. Since all three constraints in the television problem are "less-than," we respond to the questions in the following manner.

```
NUMBER OF LESS-THAN CONSTRAINTS?
?3
NUMBER OF EQUALITY CONSTRAINTS?
?0
NUMBER OF GREATER-THAN CONSTRAINTS?
?0
```

Now you will begin entering the constraints and object function formulas themselves. The program will ask you to supply the information contained in Figure 3-1 in an organized manner. It begins with the message:

```
ENTER THE COEFFICIENTS OF THE CONSTRAINTS,
                    SEPARATED BY COMMAS.
CONSTRAINT 1 ? 1, 0
```

Recall that both the controllable variables and the constraints have an order. You must enter the information using these orders. In this case, the first constraint refers to $1p + 0c \leq 60$. The coefficients are 1 and 0. Always begin with the first variable when entering coefficients, so that the coefficient of the first variable is entered first, followed by the coefficient of the second variable, and so on until the coefficients of all the variables in the constraint have been entered. LPRG knows how many coefficients to expect. If you enter too many, it will consider the last ones as extra and ignore them. You will get a message to this effect. If you do not enter enough coefficients, LPRG will print "??" until you have entered the correct number of coefficients. It will then go on to the second constraint.

```
CONSTRAINT 2 ? 0, 1
```

In this case, the second constraint is $0p + 1c \leq 40$. The coefficient of the first variable (p) is 0 and the coefficient of the second variable (c) is 1. LPRG then goes on to ask for information about the final constraint.

```
CONSTRAINT 3 ? 2, 3
```

Since the third constraint is $2p + 3c \leqslant 150$, the proper response is 2 followed by 3.

At this point, all the relevant information on the left-hand sides of the constraints has been entered.[6] LPRG knows this and will stop asking about the constraint coefficients. LPRG also knows whether the constraints involve $\leqslant$, $=$, or $\geqslant$ from the information entered earlier. The only remaining information about the constraints is the numerical values on the right-hand sides. These values are the subject of the next question.

```
ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS,
                              SEPARATED BY COMMAS.
?60,40,150
```

Remember that these values must be entered according to the order of the constraints. For this example, we respond with 60, followed by 40, followed by 150.

Now LPRG has all of the information about the constraints. Only one piece of information is missing: the coefficients of the object function.

```
ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION,
                              SEPARATED BY COMMAS.
?20,25
```

The object function is $20p + 25c = T$ so the correct response (don't forget about the order of the variables) is 20 followed by 25. Now LPRG has all the information necessary about the television problem.

Before it starts calculating, however, it gives the user a chance to correct or change the data about the constraints and the object function. If you have made a mistake in entering the data, this option allows you to correct your mistake easily by retyping the incorrect line of data.

As we have not made any mistakes in entering the data in this example, the conversation will be as follows above.

[6] In mathematical formulas, a relational operator such $\leqslant$, $\geqslant$, or = is what separates the formula into 2 sides (the right and the left). For example, in $3x + 2y + z \geqslant 0$, the left-hand side is $3x + 2y + 2$ and the right-hand side is 0. In $14a - b = 2c + d + ½$, the left-hand side is $14a$-b and the right-hand side is $2c + d + ½$.

```
DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0
```

Now LPRG will begin calculating the solution to the problem. Figure 3-1 presents a complete run of LPRG solving this problem.

3. *Interpret the results.* As you can see, the last several lines of the computer run in Figure 3-1 give the computer's answers. Let's look at them.

VARIABLE 1 refers to p and VARIABLE 2 refers to c. The maximum value of the object function, i.e., the maximum daily profit, is $1450, and will occur when $p = 60$ portables and $c = 10$ consoles. If you will refer back to page 00 you will see that these answers are the same as the ones obtained by the trial-and-error procedure.

The variables and their values will not necessarily be listed in order in the computer's answers because of the method the program uses to find the best solution. Also, sometimes a variable will not be listed under ANSWERS if it has a final value of zero. This situation is also a result of the method of solving the problem which is being used.

It is possible that there are *no* solutions to the problem you are working on. A simple example would be if we added to the constraints of the television problem that the number of portables had to be at least 50 and the number of consoles had to be at least 20. With these additional constraints the model of the problem would be

$$\left.\begin{array}{r} p \leqslant 60 \text{ (at most 60 portables)} \\ p \geqslant 50 \text{ (at least 50 portables)} \\ c \leqslant 40 \text{ (at most 40 consoles)} \\ c \geqslant 20 \text{ (at least 20 consoles)} \\ 2p + 3c \leqslant 150 \text{ (manpower limitations)} \end{array}\right\} \text{Constraints}$$

$$\left.\begin{array}{l} \text{(Maximize) } 20p + 25c \\ = T \text{ (total daily profit)} \end{array}\right\} \text{Object function}$$

This problem would have no solution because if we chose p and c to satisfy the first four con-

```
RUN
LPRG


IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?1

NUMBER OF VARIABLES?   ?2

NUMBER OF CONSTRAINTS?   ?3

NUMBER OF LESS-THAN CONSTRAINTS?   ?3

NUMBER OF EQUALITY CONSTRAINTS?   ?0

NUMBER OF GREATER-THAN CONSTRAINTS?   ?0

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1     ?1,0

CONSTRAINT  2     ?0,1

CONSTRAINT  3     ?2,3

ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?60,40,150

ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?20,25

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0


ANSWERS:
VARIABLE        VALUE
 1              60
 2              10

THE VALUE OF THE OBJECT FUNCTION IS  1450
THIS VALUE IS A MAXIMUM.

DONE
```

Figure 3-1.
Complete Run of LPRG.

straints, the last constraint would not be satisfied. For example, if we chose $p = 50$ and $c = 20$, then:

$$2 \cdot 50 + 3 + 20 = 2 \cdot 50 + 3 \cdot 20 = 160$$

which is greater than 150 and therefore violates the fifth constraint above. You might try other values for p and c in this model to convince yourself that no solution is possible.

Different linear programming programs handle this situation in different ways. Some programs give a message that there are no solutions and then quit. Others, like LPRG, will generate a solution by ignoring one or more of the constraints. LPRG is designed to test for solutions. If it finds some, it proceeds to locate the optimal solution. If LPRG cannot find any solutions, it will ignore the last constraint and test the remaining system for solutions. It will continue to eliminate constraints until it has found a problem with solutions. Then it will find the optimal solution. In the case of LPRG, therefore, it is advisable to check the solution generated against the last constraint. On occasion, you will find that the last constraint is violated by the solution. It is then up to you to decide if the solution is acceptable or if the problem must be revised in some other manner.

Also, LPRG does not distinguish those problems with an infinite number of optimal solutions from problems which have only one optimal solution. An example of a problem which has an infinite number of optimal solutions would be the television problem, assuming that the profit on consoles is $30 rather than $25. The model for the problem would then be:

$$\left.\begin{array}{r} p \leqslant 60 \\ c \leqslant 40 \\ 2p + 3c \leqslant 150 \end{array}\right\} \text{Constraints}$$

$$\text{(Maximize) } 20p + 30c = T \;\} \quad \text{Object function}$$

There would be many optimal solutions to this problem. For example, all of the following values for p and c satisfy the given constraints and produce the same maximum profit of $1500:

p	c
60	10
45	20
30	30
15	40

LPRG would give you only one of the above answers for p and c.

For any linear programming problem, there will either be no solutions, one optimal solution, or an infinite number of optimal solutions. A linear programming problem would never have, for example, exactly two optimal solutions.

REVIEW

Terminology

COEFFICIENT OF A VARIABLE: the number which multiplies the variable.

Exercises

1. In what order must the constraints of the mathematical model of a linear programming problem be listed when using LPRG? Why is this so? (Review text pages 43–44.)
2. How do you go about ordering the controllable variables? When is this order used? (Review text pages 43–45.)
3. When running LPRG, how do you tell the computer whether the object function is to be maximized or minimized? (Review text page 44.)
4. Solve the television problem yourself using LPRG, and check your answers with those in the text, page 42.)
5. Using the mathematical model you developed in Exercise 2 on pages 42–43 (the teacher aides problem), solve the problem using the computer. Compare the answers you get from the computer with the answer you obtained by trial and error.

GUIDELINES FOR FORMULATING LINEAR PROGRAMMING MODELS

PURPOSE OF THIS SECTION

The preceding section of this unit introduced the skills necessary to solve problems using linear programming. Later we will present detailed analyses of problems in educational administration which can be solved by the same means. Before analyzing problems, however, we want to discuss why this technique has been delayed in its application to education and also give a general outline for formulating linear programming models. We hope you will keep this general outline in mind as you work through the problems in the last section of this chapter. The ideas presented below should help you recognize elements and procedures common to all problems involving linear programming, even though the problems themselves deal with widely diverse problem areas.

LINEAR PROGRAMMING IN EDUCATION

Educators have only recently begun to use linear programming to solve educational problems. There are several reasons for the delay in applying the technique in this field. First, linear programming is a relatively new tool and there has not been much time for educational applications to be developed. Second, very few educators have become acquainted with this tool during their professional training programs. Third, mathematicians or computer programs needed to provide the solutions to linear programming problems have not been generally available to educators.

Probably the single most important reason for the delay in adding linear programming techniques to the battery of educational problem-solving tools has been the difficulty of expressing the goals of educational programs in terms that can be measured objectively and of relating these goals to contributing factors. Goals must be expressed objectively before linear programming techniques can be used, since the object function must be a mathematical equation. This is usually easier in business than in education because the goals of business are frequently expressed in terms of dollars or variables directly related to dollars. In the television manufacturing example you have just been working with, the goal was to maximize the daily profit, with profit expressed in terms of dollars. The relationship between the profit and the controllable variables (number of portables and number of console television sets produced) was also known. Goals involved in education often do not fall so neatly into terms that can be measured objectively. For example, how often have you seen the following educational goals expressed objectively, that is, in a quantifiable form?

Quality of teaching in a school
Student motivation
Teacher effectiveness
Suitability of a teacher for a particular course
Efficiency of maintenance scheduling in a district
Student citizenship
Student learning
Ability of students to generalize what they have learned
Social integration

Just because the above goals are seldom expressed mathematically does not mean that it is impossible to do so. The problems presented in the final section of this chapter should give you some ideas on various techniques of quantification. Of course, linear programming—and for that matter all operations research techniques—does not claim to be able to quantify *every* goal or express *every* educational problem mathematically. We are limited by the degree of development of our mathematical techniques, the presence of variables or relationships which are unknown to us, and the fact that quantifying real-life situations may require us to make so many assumptions that the final results are too far removed from the real world to be practical.

STEPS IN FORMULATING LINEAR PROGRAMMING MODELS

To help you see a common thread in all linear programming problems, and to aid you in solving linear programming problems on your own, a general list of steps in formulating models follows. The steps will be discussed in detail and illustrated by examples from the field of education and by references to the television production problem and the teacher aide problem.

Smythe and Johnson[7] list the following steps:

Steps in Formulating Linear Programming Models

1. Recognition of the problem.
2. Formulation of the mathematical model:
   a. identification of the controllable variables;
   b. choice of measure of effectiveness;
   c. mathematical representation of the object function;
   d. identification of the constraints;

[7] William R. Smythe, Jr., and Lynwood A. Johnson, *Introduction to Linear Programming, with Applications* (Englewood Cliffs, N.J.: Prentice-Hall) p. 187.

e. mathematical representation of the constraints.

Once all these steps have been accomplished, the problem is relatively easily solved by a computer program for linear programming, such as LPRG. *The major difficulty in linear programming lies in the formulation of the problem, not in the solution.*

Recognition of the Problem

In formulating the model of the problem, Step 1 is recognizing that an educational problem is appropriate for solution through linear programming. As you already know, linear programming is applicable when it is desirable to establish a program of action that is optimal in terms of the effectiveness of reaching some measurable goal and when all the familiars in the model are linear. The program of action consists of allocating or assigning some type of resource. Furthermore, this allocation or assignment must conform to certain criteria, or constraints.

In the television production problem, we were allocating television set workers and assembly lines by determining how many portable and console sets to produce in order to optimize (maximize) the daily profit. In the teacher aide problem, we were allocating jobs by determining how many full-time and part-time aides to hire in order to optimize (minimize) the amount of money required.

Other resources often allocated in educational situations include:

Teachers and teacher aides
Budget funds
Federal aid funds
Teacher time
Maintenance workers
Audio-visual equipment
School buses
Student distribution in the school system
Class time
Classrooms

What are some of the things to be optimized (either or minimized) within certain constraints in education? Consider the lists below:

Results to Maximize:
Student achievement
Teacher experience
Teacher training
Time for instruction
Availability of instructional materials
Utilization of facilities
Opportunity for extracurricular activities
Subject offerings
Nutritional value of school lunches

Results to Minimize:
Cost of:
- total education
- school lunches
- facilities
- transportation
- interest on bonds
- equipment and supplies

Pupil-teacher ratio
Transportation time
Dropouts
Distance students must travel to school
Distance students must travel between classes
Class size
Underachievement

All of the items on these lists appear to be at least potential applications of linear programming in that they are goals to be optimized.

Formulation of the Mathematical Model

The second step, 2(a), in Smythe and Johnson's procedure calls for identifying the controllable variables that affect the problem goal.

Sometimes the resources to be allocated in a linear programming problem will be the same as the controllable variables, as in the teacher aide problem, where the controllable variables were the respective numbers of full-time and part-time teacher aides to be hired. Other times, we will have to use controllable variables which are not exactly the same thing as the resources to be allocated, as in the television problem: here, the controllable variables were the number of portable and console sets to be produced, while the actual resources to be allocated were the workers and assembly lines.

In some problems, we may have to use ingenuity to express the controllable variables. Problem 3, pages 61–65, will illustrate how to use controllable variables to indicate whether or not a teacher is hired to fill a particular position.

Step 2(b) is to choose the measure of effectiveness—some criterion that can be objectively measured. In some cases this step is quite easy. Most cost objectives can be measured in terms of dollars, as the total daily profit in the television problem or the cost of hiring teacher aides. Student achievement may be measured by scores on standardized tests. Teacher experience may simply be the number of years taught. A goal such as quality of teaching, on the other hand, may be more difficult to express in objective terms, but still possible. You could devise a rating system for prospective teachers which would indicate their potential teaching quality by including such items as experience, education, recommendations, and the opinions of those who conduct teacher interviews—easily translated into numerical scores. The higher the teacher's score, the greater his potential for quality teaching. The total quality of the teachers you hire, then, would be the sum of the teacher's scores on your rating system.

Step 2(c) calls for the mathematical representation of the object function. That is, after we have decided how to objectively express the measure of effectiveness, we must write a mathematical equation for this quanity.

The measure of effectiveness for the television problem was daily profit, measured in dollars and expressed by the equation:

$$20p + 25c = T$$

where:

p = number of portables
c = number of consoles
\$20 = profit on each portable
\$25 = profit on each console
T = the measure of effectiveness = the total daily profit.

The measure of effectiveness for the teacher aide problem was the daily cost of teacher aides measured also in dollars and expressed by the equation:

$$15f + 7p = c$$

where:

f = number of full-time teacher aides
p = number of part-time teacher aides
\$15 = cost of one full-time aide per day
\$7 = cost of one part-time aide per day
C = the measure of effectiveness = the total daily cost of teacher aides

Both of the above equations are of the general form:

$$A \cdot a + B \cdot b = Z$$ [8]

where A and B are constants (numbers), a and b are controllable variables, and Z is the measure of effectiveness.

In the equation $20p + 25c = T$, the constants were $A = 20$ and $B = 25$, and we used the symbols p, c and T instead of, respectively, a, b, and Z.

In the equation $15f + 7p = C$, we had $A = 15$ and $B = 7$, and we named our controllable variables f and p. C was the symbol we used for our measure of effectiveness. Notice that an equation in this general form, which is the sum of controllable variables multiplied by constants, is a linear equation. *Recall that the object functions in linear programming problems must be linear equations.*

Of course, we are not restricted to only *two* controllable variables for the object function. If we had *five* controllable variables, the object function would be of the general form:

$$Aa + Bb + Cc + Dd + Ee = Z$$

Even when the goals of a linear programming problem can be appropriately expressed, it is often difficult to establish relationships between the goals and the controllable variables that contribute to the fulfillment of the goals. For example, we discussed above the possibility of expressing "quality of potential teaching" as the sum of the scores of the teachers who are hired. There may, however, be other variables we wish to include in calculating the total potential quality of teaching: for example, teachers' personalities will surely be a factor in quality of teaching. But how do we express mathematically the fact that hiring a certain English Department head will alienate certain other teachers in the department?

Another example of the difficulty in establishing relationships between goals and controllable variables might occur if we were measuring the quality of mathematics instruction in a school according to the amounts of time students spend in large, medium, and small groups

[8] As a convention for this chapter, numbers are represented by upper-case letters and variables by lower-case letters.

and in individual instruction. An administrator may decide that individual instruction time is generally more valuable to a student than time spent in large group instruction. The question is, then, *how much* more valuable? Twice as valuable? Ten times as valuable? Often, the effect of a controllable variable on the measure of effectiveness must be determined by a subjective judgment On the other hand, experience and research can also provide guidelines to how much effect a variable may have on a desired goal.

If the relationship between the controllable variables and the measure of effectiveness is a subjective judgment, or if research on the matter is not conclusive, linear programming is a quick, inexpensive way to try out various hypotheses and observe the resulting effects. For example, in the problem outlined above, where an administrator is interested in ascertaining how much time should be spent in various types of mathematics instruction, he or she may solve the problem several times, each time varying the relative importance of individualized instruction. The first time he or she may specify that individualized instruction is 10 times as valuable to a student as a comparable amount of large group instruction; another time, he or she may specify that those types of instruction are equally valuable. The administrator could then compare results. The results of such a simulation could be very interesting as a school district reviews or forms its philosophy regarding individualized instruction and the resulting scheduling of students.

The next step, 2(d), is to identify the constraints. They may apply to single variables or to combinations of variables. In the television problem, we had the following constraints:

No more than 60 portables could be produced daily.
No more than 40 consoles could be produced daily.
There were only 150 people available to build television sets: 3 person-days were required to produce a portable, 2 to build a console set.

In the teacher aide problem, the constraints were:

The number of full-time aides plus the number of part-time aides had to equal 85.
At least 10 full-time aides were to be hired.
At least 15 part-time aides were to be hired.
A total of at least 50 staff members were to be hired, with a full-time aide considered one staff member and a part-time aide considered one-half staff member.

Some other constraints which might typically be found in educational problems are:

There is only $50,000 in the district budget for merit increases this year.
We can only afford to pay a new science teacher $10,000.
There are only 62 seats in a school bus.
No child should spend more than one hour a day riding a school bus.
A school lunch must provide a minimum of 1,000 calories.
A teacher must have at least one free period per day.
The enrollment in drivers' education is limited to 200 students.

Step 2(e) entails mathematical representation of the constraints on the controllable variables. This procedure has been covered before (page 53). Sometimes we need to use real ingenuity to express constraints on controllable variables. For example, Problem 3 on page 61 illustrates how to express mathematically the condition that only one teacher may be hired for each available position in a school. Recall that the constraints must also be linear in order for a problem to be solvable using linear programming techniques.

Remember, [illegible] dition, that it will *not* be possible to ex[illegible] every constraint in a real problem situation mathematically. We can only hope that the simplification of the constraints imposed by the available mathematics will not alter the problem to such a degree that the results from linear programming are not useful

REVIEW

Exercise

1. List the steps required in formulating linear programming models. (Review text pages 48–51.)

USING LINEAR PROGRAMMING IN EDUCATIONAL ADMINISTRATION: TYPES OF PROBLEMS AND TECHNIQUES FOR THEIR SOLUTION

GENERAL PURPOSE OF THIS SECTION

This section is devoted to five typical problems in educational administration which may be analyzed using the operations research tool of linear programming. It may be the most important section of all, for it is here that the power and versatility of linear programming become evident. The number of different topics to which linear programming is applicable is truly remarkable, and undoubtedly more applications will occur to you as you read.

The preceding four sections outlined all the skills necessary to solve any of the following problems. They may seem more difficult than those presented up to this point *only* because there are more controllable variables in each problem and more constraints on these variables. There will always be, however, only one object function which we will want to either maximize or minimize, depending on the goal of the problem.

PROBLEM 1: SCHEDULING CLASS TIME FOR MATHEMATICS INSTRUCTION

Statement of the Problem

A school uses a flexible scheduling pattern. Mathematics classes are to be scheduled for 60 students. The schedule can include large group instruction with classes of 60, medium group instruction with classes of 30, small group instruction with classes of 15, and individual instruction with one student per teacher. In each week, there are 1,200 minutes of teacher time available, which includes preparation time as well as teaching time. It takes two minutes of preparation for each minute of large group teaching, one minute of preparation for each minute of medium group teaching, one-half minute of preparation for each minute of small group teaching, and no preparation for individual instruction. Each student must spend at least 250 minutes per week in mathematics class or individual instruction, of which at least 5 minutes must be in individual instruction and 20 minutes in small group instruction.

The school administration has placed relative values for the students on each type of instruction. The value of a unit of time of large group instruction is 2, of medium group 5, of small group 8, and of individual instruction 40. That is, based on his experience, an administrator has decided that a unit of time of individual math instruction is five times better than a unit of small group instruction, eight times better than a unit of medium group instruction, and twenty times better than a unit of large group instruction. These numbers were chosen strictly on the basis of the administrator's own judgment; another administrator might assess the values quite differently.

The problem is to determine the amount of time to be devoted to each type of instruction in order to maximize the total value of math instruction to the students.

Mathematical Model

The measure of effectiveness in this problem is the total value of mathematics instruction for one student. The total value of math instruction will depend upon the amounts of time to be devoted to each type of instruction. These amounts of time are the controllable variables. Suppose we choose the variables l, m, s, and i to represent the following values:

- l = time spent by a student in large group instruction
- m = time spent by a student in medium group instruction
- s = time spent by a student in small group instruction
- i = time spent by a student in individual instruction

The total value of math instruction will then be measured by the sum of the time spent in each type of instruction multiplied by the respective relative value for that type of instruction. From the statement of the problem, we have the following relative values for the different types of instructio

Instruction	*Relative Value*
Large group	2
Medium group	5
Small group	8
Individual	40

Therefore, the total value of mathematics instruction for each student will be:

$$2l + 5m + 8s + 40i = V$$

where we have arbitrarily chosen the variable V to represent the total value of instruction.[9] The above equation is the object function. We are interested in obtaining the highest quality of instruction possible, so our goal is to find values for l, m, s and i which will give a *maximum* value to V.

The object function we have just formed expresses the total value of mathematics instruction for each individual student. If we wished to know the value of instruction for all 60 students, we would simply multiply the equation by 60.

Equations for the constraints of the problem must now be written. That is. we must express mathematically the limitations on the controllable variables.

We first need to calculate how much teacher time will be needed for large, medium, and small group instruction and for individual instruction. A total of 1,200 minutes of teacher time is available for preparation and teaching. The class size for large group instruction is equal to the number of students taking mathematics, so just one section of large group instruction need be considered. It takes two minutes of preparation for each minute of large group instruction taught. The teacher time used for large group instruction will, therefore, be the time in preparation plus the time in teaching for one section, or $2l + l = 3l$.

The class size for medium group instruction is 30, so two sections are needed to accommodate all 60 students. It takes a minute of teacher preparation for each minute of medium group instruction taught. The teacher time used for medium group instruction will be the time in preparation plus the time in teaching multiplied by the number of sections or $2 \cdot (m + m) = 2 \cdot 2m = 4m$.

Small group teaching requires one-half minute of preparation for each minute of teaching. If s is the time for instruction for small groups, then $.5s + s = 1.5s$ is the total teacher time required for one small group. Since a small group consists of 15 students and since a total of 60 students would require four small groups, we must multiply $1.5s$ by the number of small groups required (four): $4 \times (1.5s) = 6s$. Therefore, $6s$ is the total teacher time required for four small groups.

Individual instruction requires no preparation. If one student is allotted i minutes of individual instruction, then 60 students would require $60i$ minutes of teacher time per week in individual instruction.

Now we can write the mathematical form of the constraint on teacher time: teacher time is the sum of the times spent in preparation and teaching of large, medium, and small groups and individuals and is limited to 1200 minutes per week. The mathematical statement is:

$$3l + 4m + 6s + 60i \leqslant 1200$$

The next condition we must consider is that each student must spend at least 250 minutes per week in mathematics class or individual instruction. This constraint can be expressed by:

$$l + m + s + i \geqslant 250$$ [10]

Each student must spend at least 5 minutes in individual instruction and at least 20 minutes in small group instruction. These conditions are expressed respectively by:

$$i \geqslant 5 \qquad s \geqslant 20$$

[9] All that has been done to form this equation is to *weight* the times in various types of mathematics instruction according to the assessment of the relative worth of each type of instruction.

[10] Recall that the sign $\geqslant$ means "greater than or equal to."

In summary, then, we have formed the following mathematical statements based on the original problem:

Statement in Problem	Mathematical Expression of Statement	Type of Statement.
Teacher time is limited to 1200 minutes; allow for both preparation and teaching time in all types of instruction.	$3l + 4m + 6s + 60i \leqslant 1200$	Constraint
Each student receives at least 250 minutes of math instruction.	$l + m + s + i \geqslant 250$	Constraint
Each student receives at least 5 minutes of individual instruction.	$i \geqslant 5$	Constraint
Each student receives at least 20 minutes of small group instruction.	$s \geqslant 20$	Constraint
Maximize the value of math instruction, based on relative worth of each type of instruction.	(Choose l, m, s, and i subject to the above constraints to maximize . . .) $2l + 5m + 8s + 40i = V$	Object function

We are now in a position to order the variables and the constraints and to express the constraints in a form acceptable to LPRG.

Since we have already referred to l, m, s, and i in that order, let's now formally agree that l is the first variable, m is the second, s is the third, and i is the fourth.

In assigning an order to the constraints, recall that all the "≤" constraints must come first, followed by the "=" constraints, followed by the "≥" constraints. As "$3l + 4m + 6s + 60i \leqslant 1200$" is the only "less-than" constraint, it must be first. There are no "equality" constraints; all the remaining ones are "greater-than" constraints. So we arbitrarily assign "$l + m + s + i \geqslant 250$" as the second, "$i \geqslant 5$" as the third, and "$s \geqslant 20$" as the fourth.

Since every constraint must mention each variable, it is necessary to rewrite the third and fourth constraints, using coefficients of zero for the missing variables.

$$i \geqslant 5 \text{ becomes } 0 \cdot l + 0 \cdot m + 0 \cdot s + 1 \cdot i \geqslant 5$$

$$s \geqslant 20 \text{ becomes } 0 \cdot l + 0 \cdot m + 1 \cdot i + 0 \cdot s \geqslant 20$$

Table 3-2 summarizes the problem in a form which is usable for LPRG.

	1st var l	2nd var m	3rd var s	4th var i	sign	right-hand side
1st constraint	3	4	6	60	≤	1200
2nd constraint	1	1	1	1	≥	250
3rd constraint	0	0	0	1	≥	5
4th constraint	0	0	1	0	≥	20
Object function	2	5	8	40		

Table 3-2.
Summary of school schedule problem.

REVIEW

Exercise

1 Solve the class scheduling problem presented above using LPRG, then verify your answers with those given on the following pages. Reorder the variables and run LPRG again. Are the results the same?

Analysis of the Solution

The computer program will print out the following optimal solution to the class scheduling problem. (See Figure 3-2 for the complete listing of the computer solution to this problem.)

```
GET-LPRG
RUN
LPRG


IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?1

NUMBER OF VARIABLES?   ?4

NUMBER OF CONSTRAINTS?   ?4

NUMBER OF LESS-THAN CONSTRAINTS?   ?1

NUMBER OF EQUALITY CONSTRAINTS?   ?0

NUMBER OF GREATER-THAN CONSTRAINTS?   ?3

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1    ?3,4,6,60

CONSTRAINT  2    ?1,1,1,1

CONSTRAINT  3    ?0,0,0,1

CONSTRAINT  4    ?0,0,1,0

ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?1200,250,5,20

ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?2,5,8,40

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0


ANSWERS:
VARIABLE        VALUE
 1               120
 2               105
 3               20
 4               5

THE VALUE OF THE OBJECT FUNCTION IS  1125
THIS VALUE IS A MAXIMUM.

DONE
```

Figure 3-2.
Class Scheduling Problem Using LPRG.

```
ANSWERS
THE MAXIMUM VALUE OF THE OBJECT FUNCTION IS 1125
THIS OCCURS WHEN:
  VARIABLE 2 = 105
  VARIABLE 1 = 120
  VARIABLE 3 = 20
  VARIABLE 4 = 5
ANY VARIABLES NOT LISTED HAVE 0
DONE
```

The computer prints the maximum (or minimum) value of the object function and the corresponding values of the variables, subject to the given constraints. We see that the maximum value of the object function is 1125, which occurs when variable 2 (which is m, the time spent in medium-sized group instruction) is 105 (minutes); when variable 1 (l, time spent in large group instruction) is 120; when variable 3 (s, time spent in small group instruction) is 20; and when variable 4 (i, time spent in individualized instruction) is 5. In other words, the maximum value of instruction will occur when each student receives 120 minutes of large group instruction, 105 minutes of medium group instruction, 20 minutes of small group instruction, and 5 minutes of individualized instruction. One can verify that the original conditions of the problem are satisfied by substituting the following values for the variables in the original constraints:

Optimal solution: $l = 120, m = 105, s = 20, i = 5$

First constraint: $3l + 4m + 6s + 60i = 3 \cdot 120 + 4 \cdot 105 + 6 \cdot 20 + 60 \cdot 5 = 360 + 420 + 120 + 300 = 1200 \leqslant 1200$

Second constraint: $l + m + s + i = 120 + 105 + 20 + 5 = 250 \geqslant 250$

Third constraint: $i = 5 \geqslant 5$

Fourth constraint: $s = 20 \geqslant 20$

Checking the value of the object function, we have:

$$2l + 5m + 8s + 40i$$
$$= 2 \cdot 120 + 5 \cdot 105 + 8 \cdot 20 + 40 \cdot 5$$
$$= 240 + 525 + 160 + 200$$
$$= 1125$$

REVIEW

Exercise.

1. One of the constraints in the class scheduling problem above was based on a subjective assumption by an administrator regarding the relative worth of the different types of math instruction (large, medium, and small groups and individual instruction). Using only your own intuition, try to predict how the results of the above problem will change if the administrator decides that all types of math instruction are of equal worth. What part of the math model would have to be changed? After you have made your prediction, continue reading the text.

Analysis of Exercise

If it is assumed that all types of math instruction are of equal worth, the only change that must be made in the model for the class scheduling problem is in the object function. Instead of the object function $2l + 5m + 8s + 40i = V$, we would have the new object function:

$$l + m + s + i = V$$

Since the different types of math instruction are now assumed to be equal, the times spent in each method of instruction are weighted equally (all times have a coefficient of 1).

Let us compare the answers to the class scheduling problem using this new object function with the previous answers we obtained. (See Figure 3-3 for the computer listing for the revised class scheduling problem.)

It should not be surprising that when all types of instruction are assumed to be of equal worth, the time for large group instruction is increased. Large group instruction is, after all, the most efficient form of instruction, in that all students may be accommodated in the structure of one large class taught by one teacher. The large class situation can give a large unit of

Assumption About Types of Math Instruction	Time Spent: Large Group l	Medium Group m	Small Group s	Individual Instruction i	Value of Object Function V
Weighted values: Individual instruction much more valuable	120	105	20	5	1125
All types of instruction equal	260	0	20	5	285

Table 3-3.
Class schedule problem using different weights.

```
GET-LPRG
RUN
LPRG


IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?1

NUMBER OF VARIABLES?   ?4

NUMBER OF CONSTRAINTS?   ?4

NUMBER OF LESS-THAN CONSTRAINTS?   ?1

NUMBER OF EQUALITY CONSTRAINTS?   ?0

NUMBER OF GREATER-THAN CONSTRAINTS?   ?3

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1     ?3,4,6,60

CONSTRAINT  2     ?1,1,1,1

CONSTRAINT  3     ?0,0,0,1

CONSTRAINT  4     ?0,0,1,0

ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?1200,250,5,20


ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?1,1,1,1

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0


ANSWERS:
VARIABLE        VALUE
 1               260
 3               20
 4               5

THE VALUE OF THE OBJECT FUNCTION IS  285
THIS VALUE IS A MAXIMUM.


DONE
```

Figure 3-3.
Revised Class Scheduling Problem Solved by LPRG.

time for student instruction while still keeping teacher time low.

PROBLEM 2. PLANNING LOW-COST LUNCHES

Statement of the Problem

Consider the problem of planning a school lunch menu. Assume that minimum nutritional requirements for school lunches have been established. The requirements may be stated in terms of a day, a week, or a longer period of time, as long as they are all stated in the same unit of time. A variety of foods is available. The price of each food and the nutritional value of each are known. The problem is to minimize the cost of school lunches, subject to the nutritional constraints.

Specifically, let us suppose that we wish to supply six different nutrients to school children using five types of vegetables: peas, lima beans, carrots, spinach, and beets. The nutrients are calcium, iron, phosphorus, potassium, sodium, and Vitamin C. We will designate the pounds of each vegetable purchased for every 100 children each week by the variables p (peas), l (lima beans), c (carrots), s (spinach), and b (beets). The cost per pound for each vegetable is known, and the number of grams of nutrient contained in each pound of vegetable is also known. In summary, suppose we have the information given in Table 3-4.

Vegetable	Pounds Purchased	Cost per Pound	Grams of Nutrient Per Pound of Vegetable: Calcium	Iron	Phosphorus	Potassium	Sodium	Vitamin C
Peas	p	.25	.09	.01	.26	.44	1.07	.04
Lima Beans	l	.30	.12	.01	.30	1.01	1.07	.03
Carrots	c	.19	.15	.00	.14	1.01	.15	.03
Spinach	s	.29	.42	.01	.17	1.47	.23	.13
Beets	b	.15	.06	.00	.77	.76	1.07	.01

Table 3-4.
Information for school lunch menu problem.

Furthermore, we wish to provide every 100 children with the minimum amounts of each nutrient each week given in the following list.

Nutrient	Minimum amount, in grams, required each week per 100 children
n_1	400g (100% RDA)
n_2	9g (100% RDA)
n_3	600g (100% RDA)
n_4	975g (average intake, no RDA established)
n_5	1150g (average intake, no RDA established)
n_6	20g (100% RDA)

How many pounds of each type of vegetable should be purchased each week, per 100 children, so that minimum nutritional requirements are met while at the same time spending the least amount of money possible?

Mathematical Model

The measure of effectiveness for this problem is the total cost of vegetables per week for 100 children. The controllable variables are the amounts of each type of vegetable that can be purchased.

Using the cost per pound for each vegetable and the number of pounds that must be purchased, the total cost for vegetables for 100 children for a week will be:

$$.25p + .30l + .19c + .29s + .15b = C$$

The above equation is our object function, and the total cost, C, is the quantity we wish to make as small as possible. In this problem, then, we are interested in *minimizing* the object function.

On the other hand, we wish to supply the required amounts of nutrients to the children from these foods. The minimum amount of calcium that is needed is 400 grams. For every pound of peas, .09 grams of calcium are supplied; every pound of beans provides .12 grams of calcium; and so on. Therefore, for the purchase of p, l, c, s, and b pounds of peas, lima beans, carrots, spinach and beets respectively, $(.09p + .12l + .15c + .42s + .06b)$ grams of calcium will be provided. We want this amount to be at least as large as the minimum required amount, 400 grams. Therefore, the first constraint on the number of pounds of vegetables purchased is:

$$.09p + .12l + .15c + .42s + .06b > 400$$

That is, the amount of calcium supplied to 100 children every week from these vegetables is at least as large as the minimum requirement (400 grams).

Similarly, we can calculate the total number of grams of the other nutrients supplied by the purchase of these vegetables and specify that these amounts are at least as large as the minimum requirements for the nutrients. For iron, the minimum requirement is 9 grams. Purchasing p, l, c, s, and b amounts of peas, lima beans, carrots, spinach, and beets will provide $.01p + .01l + 0c + .01s + 0b$ grams of iron. We want this amount to be at least as large as 9 grams. Therefore, our second constraint would be expressed:

$$.01p + .01l + 0c + .01s + 0b \geqslant 9$$

The requirement for phosphorus is at least 600 grams, so the third constraint would be:

$$.26p + .3l + .14c + .17s + .77b \geqslant 600$$

The last three constraints would then describe the minimum requirements for potassium, sodium, and Vitamin C:

$$.44p + 1.01l + 1.01c + 1.47s + .76b \geqslant 975$$

$$1.07p + 1.07l + .15c + .23s + 1.07b \geqslant 1150$$

$$.04p + .03l + .03c + .13s + .01b \geqslant 20$$

In summary, then, we have the following mathematical formulation of the problem: Find values for p, l, c, s, and b which are subject to the constraints

$$.09p + .12l + .15c + .42s + .06b \geqslant 400$$

$$.01p + .01l + 0c + .01s + 0b \geqslant 9$$

$$.26p + .3l + .14c + .17s + .77b \geqslant 600$$

$$.44p + 1.01l + 1.01c + 1.47s + .76b \geqslant 975$$

$$1.07p + 1.07l + .15c + .23s + 1.07b \geqslant 1150$$

$$.04p + .03l + .03c + .13s + .01b \geqslant 20$$

and which minimize

$$.25p + .30l + .19c + .29s + .15b = C$$

Notice that we have essentially already assigned an order to the variables by the manner in which we set up the problem. That is, we have always mentioned p, l, c, s, and b in that order. So we now officially recognize p is first, l is second, c is third, s is fourth and b is fifth.

Notice also that all the constraints are "greater-than," so we must assign order arbitrarily. Let's use the order in which they were developed.

All the constraints already mention each variable, so there is no need to alter the constraints. Table 3–5 summarizes the problem in a manner readily adaptable for use with LPRG.

	1st var (p)	*2nd var (l)*	*3rd var (c)*	*4th var (s)*	*5th var (b)*	*Sign*	*Right-hand side*
1st constraint	.09	.12	.15	.42	.06	>	400
2nd constraint	.01	.01	0	.01	0	⩾	9
3rd constraint	.26	.3	.14	.17	.77	⩾	600
4th constraint	.44	1.01	1.01	1.47	.76	⩾	975
5th constraint	1.07	1.07	.15	.23	1.07	⩾	1150
6th constraint	.04	.03	.03	.13	.01	⩾	20
Object function	.25	.30	.19	.29	.15		

Table 3–5.
Summary of variable information for school menu problem.

Analysis of the Solution

The amounts of vegetables which must be purchased in a week for every 100 children in the school system such that the minimum amounts of nutrients are supplied while the cost is made as small as possible are shown in the list below. (The computer solution is shown in Figure 3-4.)

```
GET-LPRG
RUN
LPRG


IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?-1

NUMBER OF VARIABLES?   ?5

NUMBER OF CONSTRAINTS?   ?6

NUMBER OF LESS-THAN CONSTRAINTS?   ?0

NUMBER OF EQUALITY CONSTRAINTS?   ?0

NUMBER OF GREATER-THAN CONSTRAINTS?   ?6

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1    ?.09,.12,.15,.42,.06

CONSTRAINT  2    ?.01,.01,0,.01,0

CONSTRAINT  3    ?.26,.3,.14,.17,.77

CONSTRAINT  4    ?.44,1.01,1.01,1.47,.76

CONSTRAINT  5    ?1.07,1.07,.15,.23,1.07

CONSTRAINT  6    ?.04,.03,.03,.13,.01

ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?400,9,600,975,1150,20


ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?.25,.30,.19,.29,.15

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0


ANSWERS:
VARIABLE        VALUE
 1               81.8836
 4               818.117
 5               817.026

THE VALUE OF THE OBJECT FUNCTION IS  380.279
THIS VALUE IS A MINIMUM.


DONE
```

Figure 3-4.
School Lunch Problem Solved Using LPRG.

Vegetable	*Amount Purchased (pounds)*
Peas (p)	81.8836
Lima Beans (l)	0
Carrots (c)	0
Spinach (s)	818.117
Beets (b)	817.026

The total cost of the purchase will be $380.28. Furthermore, this cost is the least amount which needs to be spent on vegetables for 100 children for one week in order to supply minimum nutritional requirements. It may easily be shown that this solution satisfies the original constraints by substituting values of $p = 81.8836$, $l = 0$, c

$= 0$, $s = 818.117$, and $b = 817.026$ into the set of constraints we formulated above.

REVIEW

Exercise

1. Although the above solution is mathematically sound, it may not appeal to school children who are faced with an abundance of spinach and beets in their lunches. Use your imagination to write additional constraints so that a larger variety of vegetables will be purchased without sacrificing nutritional needs. How does the total cost of the rewritten problem compare to the minimum cost above ($380.28)?
2. Do you think it is reasonable to spend $380.28 per 100 children per week for vegetables? Realistically, what could be changed in this problem to bring the price down?

PROBLEM 3. ASSIGNING TEACHERS TO POSITIONS

Statement of the Problem

Within a school, four teaching positions are to be filled by four teachers. The teaching positions and the symbols we will use to denote them are:

A	algebra	C	calculus
G	geometry	P	physics

Each of the four teachers (denoted T1, T2, T3, and T4 is qualified to fill any of the four positions.

Based on the teachers' training, experience, and recommendations and on the opinions of the administrators and department heads who have interviewed them, each teacher has been given a score from 1 to 10 according to suitability for each of the four positions. A rank of 10 indicates that a teacher is extremely well qualified for a position; 5 indicates that a teacher has average qualifications; and 1 indicates that a teacher is not well suited at all for a position, even though he may be qualified to hold it. The respective scores are given by the matrix in Figure 3-5. For example, T1 has been given a rank of 5 (average) for C, and T3 has been given a rank of 10 (extremely well qualified) for A.

Because not many teachers have applied for positions in the school, these four must be hired to fill the four positions. Your problem is to assign teachers to positions so the school will obtain the maximum possible quality of teaching.

Teachers \ Positions	A	G	C	P
T1	4	4	5	2
T2	5	4	9	7
T3	10	9	7	6
T4	8	6	4	2

Figure 3-5.
Teacher-Position Matrix.

Mathematical Model

The quality of teaching in this problem is the measure of effectiveness and will be calculated by the sum of the teachers' ranks for the positions they are assigned. Suppose we arbitrarily make the following assignment of teachers to positions:

T1	is assigned to P	T3	is assigned to A
T2	is assigned to C	T4	is assigned to B

We can indicate this assignment by drawing circles in the appropriate positions on the original teacher-position matrix, as shown in Figure 3-6.

The total value of the quality of teachers hired will be sum of the teachers' scores for those positions:

Teachers \ Positions	A	G	C	P
T1	4	4	5	**2**
T2	5	4	**9**	7
T3	**10**	9	7	6
T4	8	**6**	4	2

Figure 3–6.
Assignment of Teachers to Positions on Matrix.

Quality of teachers = 10 + 6 + 9 + 2 = 27

Obviously, the higher the total sum, the higher the potential quality of teaching. Is there a way to assign teachers to positions so that the quality of teaching is greater than 27? In posing this question, it is now evident that the problem can be solved by linear programming. We have defined quality of teaching as our measure of effectiveness, and have also defined a means of numerically calculating quality of teaching. We want to assign teachers to positions in order to *maximize* the quality of teaching.

In order to form the mathematical model for this problem, we must define our controllable variables so that we can numerically indicate whether or not a teacher is assigned to a certain position. We will need 16 controllable variables, named a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p. Each variable will be associated with the possible assignment of a teacher to a position, as indicated in the matrix below:

Teachers \ Positions	A	G	C	P
T1	a	b	c	d
T2	e	f	g	h
T3	i	j	k	l
T4	m	n	o	p

Figure 3–7.
Variables Shown on Teacher-Position Matrix.

These variables will all be used to answer the question, "Is this teacher hired for this position?" If the variable has value 0, the answer will be No. If the value is 1, the answer will be Yes. For example, we assumed arbitrarily that T1 is hired for P. Then the following values would be assigned to a, b, c, and d:

T1 assigned to P:
$a = 0$ (not assigned to A)
$b = 0$ (not assigned to G)
$c = 0$ (not assigned to C)
$d = 1$ (*yes*, assigned to P)

We also assumed T2 is assigned to C. Then the following values for e, f, g, and h would occur:

T2 assigned to C:
$e = 0$ (not assigned to A)
$f = 0$ (not assigned to G)
$g = 1$ (*yes*, assigned to C)
$h = 0$ (not assigned to P)

We assumed also that T3 is assigned to A and T4 is assigned to G, so let's complete the entire set of values for the controllable variables and compare these values to the teacher-position matrix where we circled scores for this arbitrary assignment of teachers to positions:

Teachers \ Positions	A	G	C	P
T1	4	4	5	**2**
T2	5	4	**9**	7
T3	**10**	9	7	6
T4	8	**6**	4	2

Teachers \ Positions	A	G	C	P
T1	a=0	b=0	c=0	d=1
T2	e=0	f=0	g=1	h=0
T3	i=1	j=0	k=0	l=0
T4	m=0	n=1	o=0	p=0

Figure 3–8.
Comparions of Teacher-Position Matrix and Variable Values Matrix.

Notice that there are values of 1 wherever a teacher has been assigned to a position.

Now that the controllable variables have been defined, it is easy to express the object function for the quality of teaching (denoted by Q). We take each teacher's score for each position, multiply it by the appropriate controllable variable, and add all these multiples together:

$$\begin{aligned}\text{Quality of teaching} &= 4a + 4b + 5c + 2d + 5e \\ &\quad + 4f + 9g + 7h + 10i \\ &\quad + 9j + 7k + 6l + 8m \\ &\quad + 6n + 4o + 2p \\ &= Q\end{aligned}$$

For the arbitrary assignment above, we have, substituting values for the variables:

$$\begin{aligned}\text{Quality of teaching} &= 4 \cdot 0 + 4 \cdot 0 + 5 \cdot 0 \\ &\quad + 4 \cdot 0 + 9 \cdot 1 + 7 \cdot 0 \\ &\quad + 10 \cdot 1 + 9 \cdot 0 + 7 \cdot 0 \\ &\quad + 6 \cdot 0 + 8 \cdot 0 + 6 \cdot 1 \\ &\quad + 4 \cdot 0 + 2 \cdot 0 \\ &= 2 + 9 + 10 + 6 \\ &= 27 = Q\end{aligned}$$

which is the same number we obtained before. To solve a linear programming problem on the computer, we need a mathematical expression for the object function. Defining controllable variables which could either equal 1 or 0 enabled us to do this.

Now that we have developed a mathematical expression for quality of teaching, what are the constraints on the controllable variables? Let's again examine the teacher-position matrix showing the values of the controllable variables for our arbitrary assignment.

Positions

Teachers	A	G	C	P	Sum of the values for each row
T1	a 0	b 0	c 0	d 1	1
T2	e 0	f 0	g 1	h 0	1
T3	i 1	j 0	k 0	l 0	1
T4	m 0	n 1	o 0	p 0	1
Sum of the values for each column	1	1	1	1	

Figure 3-9.
Matrix with Sum of Values.

Notice that we have indicated the sum of the values of the controllable variables for each row and each column. The fact that the sum of the values of the controllable variables for each *row* is 1 reflects the situation that only *one* controllable variable in each row has value 1; the rest of the controllable variables in that row are 0. These values of course indicate that a teacher can only be assigned to one position; if more than one controllable variable in a row equaled 1, the variables would be indicating that a teacher was assigned to more than one position—a situation not permitted. We can be assured that a teacher will be assigned to exactly one position if we specify mathematically that the sum of the controllab' 'riables in each row must equal 1. We will nee. .ne following four equations.

Condition	*Mathematical Equation*
T1 can be assigned only one position	$a + b + c + d = 1$
T2 can be assigned only one position	$e + f + g + h = 1$
T3 can be assigned only one position	$i + j + k + l = 1$
T4 can be assigned only one position	$m + n + o + p = 1$

Note: Mathematically these equations by themselves guarantee only that 100 percent of each teacher's time will be divided between the four positions and not that each teacher will spend 100 percent of his time in the same position. For instance, T1 could spend half his time on algebra and the other half on geometry and the first equation would be satisfied (½ + ½ + 0 + 0 = 1). These equations guarantee that one teacher spends all of his time on the same position *only* when we add the extra condition that the variables may only have values of 0 or 1.

For each *column* in the teacher-position matrix above, the sum of the values of the controllable variables for this arbitrary assignment of teachers is also 1. The reason why exactly one controllable variable in each column has value 1 is that each position is assigned to only one teacher. Having more than one variable equal to 1 in a column would indicate that a certain position has been assigned to more than one teacher. We therefore want to specify mathematically that each position will be assigned to exactly one teacher. We do this by stating that the sum of the values of the controllable variables for each column must be 1. Four equations are necessary to express this constraint:

Condition	Mathematical Equation
A can be assigned to only one teacher	$a + e + i + m = 1$
G can be assigned to only one teacher	$b + f + j + n = 1$
C can be assigned to only one teacher	$c + g + k + o = 1$
P can be assigned to only one teacher	$d + h + l + p = 1$

Again, notice that these equations guarantee each position will be assigned to only one teacher if it is stipulated that all the variables can have values of 0 and 1 only.

In summary, here is the mathematical model for this problem. We wish to find values for a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p such that the following constraints are satisfied:

$$\left.\begin{aligned} a+b+c+d &= 1\\ e+f+g+h &= 1\\ i+j+k+l &= 1\\ m+n+o+p &= 1 \end{aligned}\right\}\ \text{Each teacher assigned to one teaching position}$$

$$\left.\begin{aligned} a+e+i+m &= 1\\ b+f+h+n &= 1\\ c+g+k+o &= 1\\ d+h+l+p &= 1 \end{aligned}\right\}\ \text{Each teaching position has only one teacher}$$

and such that the value for Q is a maximum:

$$4a + 4b + 5c + 2d + 5e + 4f + 9g + 7h + 10i + 9j + 7k + 6l + 8m + 6n + 4o + 2p = Q$$

It is time to prepare the model for input to LPRG. Let's agree to use alphabetical order for the variables. Since all the constraints are "equality," we will number them in the order in which they were developed. Notice, however, that none of the constraints mentions all 16 variables. This means that all of them will have to be rewritten using coefficients of zero for all variables not originally mentioned in the constraint. Table 3-6 contains a summary of this problem for use with LPRG.

Note that nowhere in the model have we explicitly stated that each variable must have values of either 0 or 1. The reason for this is that such statements are not linear formulas and therefore this problem cannot be solved using linear programming techniques if they are included. We will have to be content with the eight constraints we have and be prepared to deal with the solution LPRG generates.

Analysis of the Solution

The computer listing for the solution of the teacher assignment problem is shown in Figure 3-11. The resulting values for the controllable variables are:

$$c = h = j = m = 1$$

$$a = b = d = e = f = g = i = k = l = n = o = p = 0$$

which correspond to the assignment:

T1 to calculus (C)
T2 to physics (P)
T3 to geometry (G)
T4 to algebra (A)

Positions

Teachers	A	G	C	P
T1	4	4	**5**	2
T2	5	4	9	**7**
T3	10	**9**	7	6
T4	**8**	6	4	2

Positions

Teachers	A	G	C	P
T1	a=0	b=0	c=1	d=0
T2	e=0	f=0	g=0	h=1
T3	i=0	j=1	k=0	l=0
T4	m=1	n=0	o=0	p=0

Figure 3-10.
Teacher-Position and Variable-Value Matrices for Problem Solution.

The resulting quality of teaching, measured by the sum of the teachers' scores for the assigned positions, is the maximum possible figure:

Quality of teaching = 5 + 7 + 9 + 8 = 29 = Q

Linear programming can be quite useful to administrators in solving such assignment problems as the one outlined above. This problem could probably be solved by trial and error in not too long a time, since there are only 24 ways to assign four teachers to four positions. However, consider the situation of assigning *ten* teachers to *ten* positions: there are over *3½ million* ways to make these assignments! It

	1st var (a)	2nd var (b)	3rd var (c)	4th var (d)	5th var (e)	6th var (f)	7th var (g)	8th var (h)	9th var (i)	10th var (j)	11th var (k)	12th var (l)	13th var (m)	14th var (n)	15th var (o)	16th var (p)	Sign	Right-hand side
1st constraint	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	=	1
2nd constraint	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	=	1
3rd constraint	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	=	1
4th constraint	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	=	1
5th constraint	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	=	1
6th constraint	0	1	0	0	0	1	0	0	[illegible]	1	0	0	0	1	0	0	=	1
7th constraint	0	0	1	0	0	[illegible]	1	0	0	0	1	0	0	0	1	0	=	1
8th constraint	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	=	1
Object function	4	4	5	2	5	4	9	7	10	9	7	6	8	6	4	[illegible]		

Table 3-6.
Summary of teacher assignment problem.

would be virtually impossible to manually examine those 3½ million assignments in order to find the optimal assignment. Such a problem could, however, easily be solved in a few minutes using linear programming and LPRG.

A FOOTNOTE ON ASSIGNMENT PROBLEMS

The technique of defining controllable variables so that they indicate the presence or absence of a condition (such as whether or not a teacher is assigned to a certain position) is fairly common in operations research procedures and is useful in a wide variety of problems. For example, variables could have value 1 or 0 according to whether or not a school bus travels a specific portion of a route, or whether or not a student is assigned to a first-period English class.

Here are some variations of assignment problems which can be solved using techniques similar to those above:

- assigning students to classes;
- making up a master schedule for a high school;
- assigning topics to be covered to teachers who are team-teaching a course;
- scheduling use of maintenance equipment for a large school system.

In some of the assignment problems suggested above, it might be more appropriate to use numerical ranks according to *preferences* instead of using scores based on ability, as we did in the teacher assignment problem. For example, a group of teachers who are team-teaching a course might numerically rank the topics in the course according to their personal preferences for teaching each of the topics; the higher the number for a topic, the greater the preference of a teacher for teaching it. An object function could then be formed to represent the total value of teachers' preferences (we might call this the "happiness total"). If teachers were assigned to teach topics in the course so that this object function was a *maximum*, we would be making the assignment which resulted in the greatest possible satisfaction among the teachers.

PROBLEM 4. BUSING STUDENTS TO ACHIEVE RACIAL BALANCE

Statement of the Problem

An inner-city area has nine designated school bus stops of various racial mixes which are to feed two newly constructed elementary schools, Edgar Allen Poe Elementary (P) and Robert Frost Elementary (F). Each school has room for not more than 400 students.

```
GET-LPRG
RUN
LPRG
IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?1

NUMBER OF VARIABLES?   ?16

NUMBER OF CONSTRAINTS?   ?8

NUMBER OF LESS-THAN CONSTRAINTS?   ?0

NUMBER OF EQUALITY CONSTRAINTS?   ?8

NUMBER OF GREATER-THAN CONSTRAINTS?

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1    ?1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0

CONSTRAINT  2    ?0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,0

CONSTRAINT  3    ?0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0

CONSTRAINT  4    ?0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1

CONSTRAINT  5    ?1,0,0,0,1,0  ,0,1,0,0,0,1,0,0,0

CONSTRAINT  6    ?0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0

CONSTRAINT  7    ?0,0,1,0,0,0,1,0,0,0,1,0,0,0,1,0

CONSTRAINT  8    ?0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,1
ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?1,1,1,1,1,1,1,1

ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?4,4,5,2,5,4,9,7,10,9,7,6,8,6,4,2

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0
(

ANSWERS:
VARIABLE        VALUE
 3               1
 7               0
 8               1
 9               0
 10              1
 12
 13              1

THE VALUE OF THE OBJECT FUNCTION IS   29
THIS VALUE IS A MAXIMUM.

DONE
```

Figure 3-11.
Teacher Assignment Problem Solved Using LPRG.

The school board has determined that all the students at the same bus stop will go to the same school, that each of the schools must have a minimum of 180 black students and 130 white students, and that all students should be bused as short a distance as possible. The distance traveled will be measured in student-miles, which is the total number of miles traveled by all students (for example, if one student is bused 3 miles and another 7 miles, the total number of student-miles traveled is 10).

The total number of students per stop, the number of black students and white students at each stop, and the distance of each stop from each school are summarized in Table 3-7.

Let's arbitrarily assign bus stops 1, 2, 3, 4, and 5 to go to P and bus stops 6, 7, 8, and 9 to go to F and use Table 3-7 to compute the num-

Stop	Total # Students	# Black Students	# White Students	Distance to P	Distance to F
1	100	100	0	.2	5.0
2	90	80	10	1.0	4.5
3	80	65	15	2.0	4.0
4	75	55	20	1.0	4.5
5	80	40	40	2.5	2.5
6	85	30	55	2.5	2.0
7	80	35	45	3.0	2.0
8	70	10	60	4.0	1.5
9	50	0	50	5.0	.3

Table 3-7.
Summary of bus stop data.

ber of student miles traveled. There are 100 students at bus stop 1, .2 miles from P, so the total number of student-miles from bus stop 1 is 100 × .2 = 20 miles. There are 90 students at bus stop 2, 1 mile from P, making the total number of student-miles from bus stop 2 90 × 1 = 90 miles. Similarly, bus stop 3 accounts for 80 × 2 = 160 student-miles, bus stop 4 for 75 × 1 = 75 student-miles, and bus stop 5 for 80 × 2.5 = 200 student-miles. Therefore, the total number of student miles traveled by students going to P is:

Bus stop 1 Bus stop 2 Bus stop 3
(100 × .2) + (90 × 1) + (80 × 2)
Bus stop 4 Bus stop 5
+ (75 × 1) + (80 × 2.5) = 545 student-miles.

In the same manner we calculate the number of student-miles traveled by students going to F as:

Bus stop 6 Bus stop 7 Bus stop 8
(85 × 2) + (80 × 2) + (70 × 1.5)
Bus stop 9
+ (50 × .3) = 170 + 160 + 105 + 15 = 450
student-miles.

Thus the total number of student miles traveled is

No. Student-miles traveled by students of P
+ Student-miles traveled by students of F
= 545 + 450 = 995 student-miles.

Let's see if this busing assignment conforms to the constraints set. The first is that each school can accommodate no more than 400 students. The number of students going to P is the number at the first five bus stops, which, according to Table 3-7, is 100 + 90 + 80 + 75 + 80 = 425. The number of students going to F is 85 + 80 + 70 + 50 = 285. Clearly there are too many students at P; and so these constraints are not satisfied.

The next set of constraints has to do with how many black and how many white students go to each school. Again using Table 3-7, we see that P has 100 + 80 + 65 + 55 + 40 = 340 black students and 0 + 10 + 15 + 20 + 40 = 85 white students. F has 30 + 35 + 10 + 0 = 75 black students and 55 + 45 + 60 + 50 = 210 white students. Thus, P does not have enough white students—each school must have a minimum of 130— and F does not have enough black students: each school must have a minimum of [illegible]0. So these constraints are not satisfied either.

Our arbitrary assignment of a busing pattern, then, does not solve the problem because it does not satisfy the constraints. Rather than assigning another pattern and calculating all the constraints again in hopes of finding a solution and then trying to find the best solution, let's turn our effort to building a precise mathematical model for the situation, hoping to find it has all linear formulas so that we can use LPRG to solve the problem.

In order to form a mathematical model for this problem, we will need to define controllable variables having to do with each bus stop and where its students go to school. This can be accomplished by defining two variables per bus stop, one indicating whether they go to Poe, the other telling whether they go to Frost. Each variable will have a value of 1 if the students go to the school and 0 if they do not. Notice that this technique of assigning variables is similar to that in the previous problem. For the first bus stop, we will call the Poe-indicator P1 and the Frost-indicator F1; for the second, P2 and F2; and so on through the ninth bus stop, with P9 and F9. It should be obvious that for each pair of variables associated with a given bus stop, one must have a value of 1 and the other must have a value of 0. The following list shows the values of all 18 variables for the busing pattern discussed earlier—that is, that the first five bus stops go to P and the last four go to F.

Bus Stop	*P-indicator*	*F-indicator*
1	P1 = 1	F1 = 0
2	P2 = 1	F2 = 0
3	P3 = 1	F3 = 0
4	P4 = 1	F4 = 0
5	P5 = 1	F5 = 0
6	P6 = 0	F6 = 1
7	P7 = 0	F7 = 1
8	P8 = 0	F8 = 1
9	P9 = 0	F9 = 1

Now that we know what the controllable variables are, the next step is to formulate the constraints in terms of them. First, let's work on the constraints that limit each school size to a maximum of 400 students. There will actually be two constraints in this set, one for P and one for F. In order to calculate the number of students going to P, we need to add up the number of students at each bus stop to be assigned to P. But we do not know ahead of time which of the nine bus stops will be assigned to P and which of them will be assigned to F, and this presents a problem. We do have, however, a P-indicator for each bus stop. The value of the P-indicator tells us whether or not the students at that stop are going to P (the value is 1 if they are, 0 if they are not). Therefore the quantity

Total no. students at bus stop
× P-indicator for bus stop

will be the total number of students from that bus stop who are going to P. Notice that if the students at the bus stop are going to F, the P indicator will be 0 and so the whole quantity will be 0. So an expression which tells the total number of students going to P is:

Bus stop 1	Bus stop 2	Bus stop 3
(100 × P1) +	(90 × P2) +	(80 × P3)
Bus stop 4	Bus stop 5	Bus stop 6
(75 × P4) +	(80 × P5) +	(85 × P6)
Bus stop 7	Bus stop 8	Bus stop 9
(80 × P7) +	(70 × P8) +	(50 × P9)

and the constraint that P can have no more than 400 students is expressed by:

$$100P1 + 90P2 + 80P3 + 75P4 + 80P5 + 85P6 + 80P7 + 70P8 + 50P9 \leq 400$$

Similarly, the constraint that F can have no more than 400 students is expressed by

$$100F1 + 90F2 + 80F3 + 75F4 + 80F5 + 85F6 + 80F7 + 70F8 + 50F9 \leq 400$$

Going back to the busing pattern discussed earlier and using the values of the variables as defined in the list, we can evaluate the left-hand side of the P constraints as:

$$\begin{aligned} & 100P1 + 90P2 + 80P3 + 75P4 + 80P5 + 85P6 + 80P7 + 70P8 + 50P9 \\ = \; & 100 \cdot 1 + 90 \cdot 1 + 80 \cdot 1 + 75 \cdot 1 + 80 \cdot 1 + 85 \cdot 0 + 80 \cdot 0 + 70 \cdot 0 + 50 \cdot 0 \\ = \; & 100 + 90 + 75 + 80 + 0 + 0 + 0 + 0 \\ = \; & 425 \end{aligned}$$

which is the same figure we originally calculated in connection with this set of constraints.

Now let's turn our attention to the set of constraints concerning the minimum number of black and white students at each school. There will actually be four of these, one for black students at P, one for white students at P, one for black students at F, and one for white students at F.

Let's work on the number of black students at P first. This will be the sum of the number of black students at all of the bus stops assigned to P. Again, since we do not know the bus stop assignments in advance, we must use the P-indicators. The quantity

No. black students at bus stop
× P-indicator for that stop

gives the total number of black students at that bus stop who are going to P. Again, notice that if the bus stop is assigned to F, the P-indicator will have a value of 0 and the number of black students going to P from that bus stop will be 0 also. So, using Table 3-7, an expression of the total number of black students who are going to P is:

Bus stop 1	Bus stop 2	Bus stop 3
(100 × P1) +	(80 × P2) +	(65 × P3)

Bus stop 4	Bus stop 5	Bus stop 6
$+ (55 \times P4)$	$+ (40 \times P5)$	$+ (30 \times P6)$
Bus stop 7	Bus stop 8	Bus stop 9
$+ (35 \times P7)$	$+ (10 \times P8)$	$+ (0 \times P9)$

and the constraint that P must have at least 180 black students is expressed by:

$$100P1 + 80P2 + 65P3 + 55P4 + 40P5 + 30P6 + 35P7 + 10P8 + 0P9 \geq 180$$

Similarly, the constraint that F must have at least 180 black students is expressed by:

$$100F1 + 80F2 + 65F3 + 55F4 + 40F5 + 30F6 + 35F7 + 10F8 + 0F9 \geq 180$$

The constraint about P's white students is constructed by multiplying the total number of white students at each stop (from Table 3–7) by the P-indicator and then adding together the number of white students at each stop who are going to P, to get:

Bus stop 1	Bus stop 2	Bus stop 3
0P1	+ 10P2	+ 15P3
Bus stop 4	Bus stop 5	Bus stop 6
+ 20P4	+ 40P5	+ 55P6
Bus stop 7	Bus stop 8	Bus stop 9
+ 45P7	+ 60P8	+ 50P9

So, the constraint that P must have at least 130 white students is expressed by:

$$0P1 + 10P2 + 15P3 + 20P4 + 40P5 + 55P6 + 45P7 + 60P8 + 50P9 \geq 130$$

and the constraint that F must have at least 140 students is expressed by:

$$0F1 + 10F2 + 15F3 + 20F4 + 40F5 + 55F6 + 45F7 + 60F8 + 50F9 \geq 130$$

Evaluating the left-hand side of the constraint about P's black students using the variable values from the sample busing pattern discussed earlier, we have:

$$\begin{aligned} & 100P1 + 80P2 + 65P3 + 55P4 + 40P5 \\ & \quad + 30P6 + 35P7 + 10P8 + 0P9 \\ = {} & 100 \cdot 1 + 80 \cdot 1 + 65 \cdot 1 + 55 \cdot 1 + 40 \cdot 1 \\ & \quad + 30 \cdot 0 + 35 \cdot 0 + 10 \cdot 0 + 0 \cdot 0 \\ = {} & 100 + 80 + 65 + 55 + 40 \\ & \quad + 0 + 0 + 0 + 0 \\ = {} & 340 \text{ black students} \end{aligned}$$

which is the same number of black students going to P that we calculated originally.

Although it may seem that we have calculated all the constraints for this problem, we have left out a very important set of them: the artificial constraints we put on the controllable variables so that the students at each bus stop are assigned to only one school. We will need nine of these constraints, one for each pair of variables which corresponds to one for each bus stop. A simple way of expressing the constraint is:

$$\text{P-indicator} + \text{F-indicator} = 1$$

Here is a list of all nine constraints in this set:

$P1 + F1 = 1$	$P6 + F6 = 1$
$P2 + F2 = 1$	$P7 + F7 = 1$
$P3 + F3 = 1$	$P8 + F8 = 1$
$P4 + F4 = 1$	$P9 + F9 = 1$
$P5 + F5 = 1$	

Notice that we are in a situation similar to that in Problem 3, namely that we would like to stipulate that the variables can have values of 0 and 1 only. However, the nine constraints listed above do *not do this*. (For instance, the first constraint would be satisfied by $P1 = 3/4$ and $F1 = 1/4$.) These constraints do guarantee that each student at each bus stop is assigned to only one school (the same student cannot go to both schools). As in Problem 3, we will have to use these constraints, because constraints that restrict variable values to 0 and 1 are not linear formulas and so their inclusion in the model would make this problem impossible to solve using linear programming techniques.

Finally, we can turn our attention to the object function. Recall that we wish to keep the number of student-miles traveled to a minimum. Since all the students at the same bus stop are the same distance from their school, we can calculate the number of student-miles traveled by all the students at that bus stop by:

No. students at bus st[illegible]
× distance from assigned school

Again, since we do not know ahead of time whether the students at the bus stop will be going to P or to F, we will have to use the P and F indicators. Consider the following quantity:

No. students at bus stop
× distance from P × P-indicator
+ No. students at bus stop × distance from F
× F-indicator

Notice that the first part computes the number of student-miles traveled by students at the bus stop who are going to P, the second part those who are going to F. Remember that the students are going to only one school; therefore, exactly one of the school-indicators will have the value of 1, the other 0, meaning that the students will be counted only once, as going to their assigned school, whichever one that may be. Since the object function adds an expression like the one above for each bus stop, it expresses the total number of student-miles traveled by all the students to their respective schools. Using the values from Table 3-7, the object function is:

$$\begin{aligned}
100(2)P1 &+ 100(5)F1 + 90(1)P2 + 90(4.5)F2 \\
&+ 80(2)P3 + 80(4)F3 + 75(1)P4 \\
&+ 75(4.5)F4 + 80(2.5)P5 + 80(2.5)F5 \\
&+ 85(2.5)P6 + 85(2)F6 + 80(3)P7 \\
&+ 80(2)F7 + 70(4)P8 + 70(1.5)F8 \\
&+ 50(5)P9 + 50(3)F9 = M
\end{aligned}$$

or

$$\begin{aligned}
20P1 &+ 500F1 + 90P2 + 405F2 \\
&+ 160P3 + 320F3 + 75P4 \\
&+ 337.5F4 + 200P5 + 200F5 \\
&+ 212.5P6 + 170F6 + 240P7 \\
&+ 160F7 + 280P8 + 105F8 \\
&+ 250P9 + 15F9 = M
\end{aligned}$$

In summary, we have defined 18 controllable variables and a P-indicator and an F-indicator for each bus stop. We are trying to find values for the variables that will minimize the number of student-miles traveled while the following constraints are conformed to:

$$\left.\begin{aligned}
100P1 &+ 90P2 + 80P3 + 75P4 \\
&+ 80P5 + 85P6 + 80P7 \\
&+ 70P8 + 50P9 \leqslant 400 \\
100F1 &+ 90F2 + 80F3 + 75F4 \\
&+ 80F5 + 85F6 + 80F7 \\
&+ 70F8 + 50F9 \leqslant 400
\end{aligned}\right\}\ \text{Each school can have a maximum enrollment of 400 students}$$

$$\left.\begin{aligned}
100P1 &+ 80P2 + 65P3 + 55P4 \\
&+ 40P5 + 30P6 + 35P7 \\
&+ 10P8 + 0P9 \leqslant 180 \\
100F1 &+ 80F2 + 65F3 + 55F4 \\
&+ 40F5 + 30F6 + 35F7 \\
&+ 10F8 + 0F9 \leqslant 180
\end{aligned}\right\}\ \text{Each school must have at least 180 black students}$$

$$\left.\begin{aligned}
0P1 &+ 10P2 + 15P3 + 20P4 \\
&+ 40P5 + 55P6 + 45P7 \\
&+ 60P8 + 50P9 \leqslant 130 \\
0F1 &+ 10F2 + 15F3 + 20F4 \\
&+ 40F5 + 55F6 + 45F7 \\
&+ 60F8 + 50F9 \leqslant 130
\end{aligned}\right\}\ \text{Each school must have at least 130 white students}$$

$$\left.\begin{aligned}
P1 + F1 &= 1 \\
P2 + F2 &= 1 \\
P3 + F3 &= 1 \\
P4 + F4 &= 1 \\
P5 + F5 &= 1 \\
P6 + F6 &= 1 \\
P7 + F7 &= 1 \\
P8 + F8 &= 1 \\
P9 + F9 &= 1
\end{aligned}\right\}\ \text{Each student at each bus stop can go to only one school}$$

$$\left.\begin{aligned}
20P1 &+ 500F1 + 90P2 \\
&+ 405F2 + 160P3 \\
&+ 320F3 + 75P4 \\
&+ 337.5F4 + 200P5 \\
&+ 200F5 + 212.5P6 \\
&+ 170F6 + 240P7 \\
&+ 160F7 + 280P8 \\
&+ 105F8 + 250P9 \\
&+ 15F9 = M
\end{aligned}\right\}\ \text{Object Function}$$

Since we have all linear formulas in the mathematical model, we can solve this problem by using linear programming techniques with the help of the program LPRG. Before we can use LPRG, however, we must assign an order to the variables and to the constraints, and we must make sure that all the formulas are in an acceptable form.

We will agree to order the variables by bus stop, with all the P-indicators first. The order of the variables, then, is P1, P2, P3, P4, P5, P6, P7, P8, P9, F1, F2, F3, F4, F5, F6, F7, F8, F9. This is an arbitrary assignment of order which arises from the preceding discussion. Several others could have been used just as readily.

Since all the "less-than" constraints must come first, the constraints concerning the schools' maximum size must come first. We will agree that the constraint for P will be the first and for F the second (again, this internal order is arbitrary). The "equality" constraints must come next: the constraints that limit the students at each bus stop to being assigned to one school. We will agree that they will be internally ordered according to bus-stop order. Finally come the "greater-than" constraints, which con-

	Var (P1)	Var (P2)	Var (P3)	Var (P4)	Var (P5)	Var (P6)	Var (P7)	Var (P8)	Var (P9)	Var (F1)	Var (F2)	Var (F3)	Var (F4)	Var (F5)	Var (F6)	Var (F7)	Var (F8)	Var (F9)		Sign	Right-hand side
1st constraint	100	90	80	75	80	85	80	70	50	0	0	0	0	0	0	0	0	0		[illegible]	400
2nd constraint	0	0	0	0	0	0	0	0	0	100	90	80	75	80	85	80	70	50		[illegible]	400
3rd constraint	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0		[illegible]	1
4th constraint	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0		[illegible]	1
5th constraint	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0		[illegible]	1
6th constraint	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0		[illegible]	1
7th constraint	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0		[illegible]	1
8th constraint	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0		[illegible]	1
9th constraint	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0		[illegible]	1
10th constraint	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0		[illegible]	1
11th constraint	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1		[illegible]	1
12th constraint	100	80	65	55	40	30	35	10	0	0	0	0	0	0	0	0	0	0		[illegible]	180
13th constraint	0	0	0	0	0	0	0	0	0	100	80	65	55	40	30	35	10	0		[illegible]	180
14th constraint	0	10	15	20	40	55	45	60	50	0	0	0	0	0	0	0	0	0		[illegible]	140
15th constraint	0	0	0	0	0	0	0	0	0	0	10	15	20	40	55	45	60	50		>	140
Object function	20	90	160	75	200	212.5	240	280	250	500	405	320	337.5	200	170	160	105	15			

Table 3-8.
Summary of results for busing problem.

cern minimum enrollments of black and white students at each school. We will agree to order these as they were developed—that is, black students at P, black students at F, white students at P, and white students at F.

Following is a list showing all the constraints in order and the object function. Within each constraint the variables are also in order. Note the object function had to be rewritten using the agreed-on order of the variables.

$$100P1 + 90P2 + 80P3 + 75P4 + 80P5 + 85P6 + 80P7 + 70P8 + 50P9 \leqslant 400$$

$$100F1 + 90F2 + 80F3 + 75F4 + 80F5 + 85F6 + 80F7 + 70F8 + 50F9 \leqslant 400$$

$$P1 + F1 = 1 \qquad P6 + F6 = 1$$
$$P2 + F2 = 1 \qquad P7 + F7 = 1$$
$$P3 + F3 = 1 \qquad P8 + F8 = 1$$
$$P4 + F4 = 1 \qquad P9 + F9 = 1$$
$$P5 + F5 = 1$$

$$100P1 + 80P2 + 65P3 + 55P4 + 40P5 + 30P6 + 35P7 + 10P8 + 0P9 \geqslant 180$$

$$100F1 + 80F2 + 65F3 + 55F4 + 40F5 + 30F6 + 35F7 + 10F8 + 0F9 \geqslant 180$$

$$0P1 + 10P2 + 15P3 + 20P4 + 40P5 + 55P6 + 45P7 + 60P8 + 50P9 \geqslant 130$$

$$0F1 + 10F2 + 15F3 + 20F4 + 40F5 + 55F6 + 45F7 + 60F8 + 50F9 \geqslant 130$$

$$20P1 + 90P2 + 160P3 + 75P4 + 200P5 + 212.5P6 + 240P7 + 280P8 + 250P9 + 500F1 + 405F2 + 320F3 + 337.5F4 + 200F5 + 170F6 + 160F7 + 105F8 + 15F9 = M$$

One last step remains before we are ready to use LPRG on this problem: rewriting the constraints so that all the variables are mentioned. Of course, whenever new variables are added to an existing formula the corresponding coefficient is 0. Table 3-8 summarizes the constraints and the object function by listing the coefficient of each variable, the sign of the formula, and the value on the right-hand side.

Analysis of the Solution

The LPRG run for this problem is Figure 3-12 on page 72. The results are summarized in Table 3-9. Note that all the variables have values of 0 or 1 except for the second and eleventh variables, P2 and F2. Since this does not meet the school board's specifications, we will have to assess this solution further.

We have a choice how to proceed. As a first strategy, we might accept the solution generated by LPRG and split the students at the second bus stop between the two schools. This would mean making two trips to the second bus stop, one to pick up the Poe students and one to pick up the Frost students. This also goes against the school boards' specification that all the students from a bus stop go to the same school. As a second strategy, we might decide arbitrarily to assign all the students at the second bus stop to one or the other school. This would satisfy the school board in that each bus stop would then be assigned to only one school; however, if we assign the students to P, P will have 420 students assigned when it can hold

Bus Stop	P-indicator	F-indicator	School to Which Students are Assigned
1	1	0	Poe
2	½	½	half to Poe, half to Frost
3	0	1	Frost
4	1	0	Poe
5	0	1	Frost
6	1	0	Poe
7	0	1	Frost
8	1	0	Poe
9	0	1	Frost

Table 3-9.
Summary of Results for Busing Problem.

```
GET-LPRG
RUN
LPRG

IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?-1

NUMBER OF VARIABLES?   ?18

NUMBER OF CONSTRAINTS?   ?15

NUMBER OF LESS-THAN CONSTRAINTS?   ?2

NUMBER OF EQUALITY CONSTRAINTS?   ?9

NUMBER OF GREATER-THAN CONSTRAINTS?   ?4

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1    ?100,90,80,75,80,85,80,70,50,0,0,0,0,0,0,0,0,0

CONSTRAINT  2    ?0,0,0,0,0,0,0,0,0,100,90,80,75,80,85,80,70,50

CONSTRAINT  3    ?1,0,0,0,0,0,0,0,0,1,0,0,[illegible],0,0,0,0

CONSTRAINT  4    ?0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0

CONSTRAINT  5    ?0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0

CONSTRAINT  6    ?0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0

CONSTRAINT  7    ?0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0

CONSTRAINT  8    ?0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0

CONSTRAINT  9    ?0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0

CONSTRAINT  10   ?0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0

CONSTRAINT  11   ?0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1

CONSTRAINT  12   ?100,80,65,55,40,30,35,10,0,0,0,0,0,0,0,0,0,0

CONSTRAINT  13   ?0,0,0,0,0,0,0,0,0,100,80,65,55,40,30,35,10,0

CONSTRAINT  14   ?0,10,15,20,40,55,45,60,50,0,0,0,0,0,0,0,0,0

CONSTRAINT  15   ?0,0,0,0,0,0,0,0,0,0,10,15,20,40,55,45,60,50

ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?400,400,1,1,1,1,1,1,1,1,1,180,180,140,140

ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?20,90,160,75,200,212.5,240,280,250,500,405,320,337.5,200,170,160,105,15

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0
```

Figure 3-12.
Busing Problem Solved Using LPRG.

```
DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0

ANSWERS:
VARIABLE        VALUE
 1              1.
 2              .5
 4              1
 5              0
 6              1
 8              1.
 11             .5
 12             1
 14             1.
 16             1
 18             1

THE VALUE OF THE OBJECT FUNCTION IS  1530.
THIS VALUE IS A MINIMUM.

DONE
```

Figure 3-12 (Continued)

only 400, and F will have only 140 black students when it should have at least 180. On the other hand, if we assign the students at the second bus stop to F, then P will only have 135 white students when it should have at least 140. In either case, the value of the object function will be changed. If the students go to P it will be less (1372.5), and if they go to F it will be more (1687.5). A third possible strategy is to choose a solution that assigns all the students at a bus stop to the same school. But this would mean that the value of the object function is larger than the minimum of 1530.

Our inability to state the constraints of this problem completely in linear formulas has resulted in an optimal solution which is not completely satisfactory. It must be recognized that the fault is not in LPRG or even in the technique of linear programming; the trouble is that we have taken a basic problem which can be solved by *linear* programming techniques and added *nonlinear* constraints to it. We then ignored the nonlinear constraints, used LPRG to solve the linear portion of the problem, and hoped that the solution would also conform to our extra constraints. This happened in Problem 3; but in Problem 4 it did not work out so nicely. Since we created this dilemma artificially, it is now up to us to decide which of the constraints must be kept and which may be relaxed and/or ignored.

REVIEW

Exercises

1. Which of the strategies outlined above would you choose to adopt for dealing with the busing problem? Give your reasons.
2. Many factors other than mileage must be considered when determining school bus routes in real-world situations. Using your own experience,
   (a) List as many criteria as you can which could be controllable variables in an actual bus routing problem (for example, the length of time a child must spend on the bus).
   (b) Give as many realistic constraints as you can for the controllable variables you listed in (a) above. For example: no child should have to ride a bus for more than one hour each day.
   (c) Give as many examples as you can of different measures of effectiveness which would be appropriate for a bus routing problem—for example, the total amount of time all school children spend riding buses should be a minimum.

PROBLEM 5. DETERMINING A SALARY SCHEDULE

Background

One of the most complex and time-consuming tasks of the public school administrator is constructing salary schedules. Teachers' unions, support staff organizations, school board, and taxpayers all put pressure on the school district administration to develop an equitable salary schedule; and that difficult task must be accomplished with a set amount of funds.

Most school districts operate on a fixed-step salary schedule. This means that there are separate schedules for teachers and administrators, usually based only on the criteria of advanced degrees and years of experience. Salary schedule revisions typically raise questions not easily resolved using such a system, for example:

Should the district provide financial incentive for teachers to obtain college credits beyond their degree? If so, how should this incentive be distributed?

How should teachers' salaries compare to administrator's salaries?

How should the salary schedule allow compensation for the teacher who is also a department head? Should all department heads be on a separate schedule from the teachers?

Should a teacher or administrator in an inner-city school receive a comparatively higher salary than a teacher or administrator in a suburban school? If so, how could we build this condition into the schedule?

Should teaching excellence be financially rewarded? If so, how?

How can we revise the salary schedule so that younger teachers are attracted to the district?

James E. Bruno has listed a number of criteria which should be part of an effective evaluation scheme for determining salaries in a school district. He summarizes these criteria by stating:

> An "ideal" salary schedule should be able to consider all factors which both teacher groups and school boards consider important in the evaluation of salary and to the attainment of the school district objectives. Most important, the salary evaluation scheme should be internally consistent and, if desired, maintain the organizational salary hierarchy. It should be flexible enough, however, to permit overlaps in salary between the various hierarchies (e.g., it should permit a highly qualified teacher to receive more salary than a low qualified administrator). Finally, an effective salary schedule should be able to consider the financial constraints placed upon school district resources available to support the salary structure. This latter feature could be extremely important in collective bargaining negotiations between teacher unions and school boards.[11]

Bruno contends that the traditional fixed-step salary schedule does not fulfill these criteria. As an alternative, he has developed a mathematical model based on the technique of linear programming. We will present a version of Bruno's model that simplifies the mathematical notation, eliminates or revises some of the constraints, and uses only five evaluation factors in determining salaries instead of Bruno's nine. After you have worked through the problem, you may want to read Bruno's complete article.

Statement of the Problem

You have just completed conducting salary negotiation [illegible] the teachers and school board in your di[illegible] and the following general criteria have been agreed upon by all groups in determining a new salary schedule:

1. Five "district functions" are recognized: Superintendent (and Assistants to the Superintendent), Administrator (Principals, Assistant Principals, Vice-Principals), Department Head, Teacher, and Teacher Aide. The lowest possible salary a person can earn in each of these functions is as follows:

Superintendent	$22,000
Administrator	15,800
Department Head	11,500
Teacher	8,500
Teacher Aide	6,000

That is, a teacher with minimum qualifications (B.A. or B.S. degree and no experience) would receive $8500; the lowest

[11] James E. Bruno, "An Alternative to the Fixed Step Salary Schedule," *Educational Administration Quarterly*, 6 (1970), 26–46; quotation, pp. 29–30.

paid Administrator would receive $15,800; and so on.

2. Five evaluation factors will determine the salaries of teachers and administ tors within the district. Each factor will determine part of each teacher's or administrator's salary; that is, a teacher or administrator will be assessed within each factor and a certain amount of compensation assigned for that factor. The total salary of a teacher or administrator will then be the sum of the amounts from each of the five evaluation factors. The evaluation factors are:
   a. *Learning environment.* Those personnel with "difficult" assignments, such as inner-city schools, will receive more compensation from this factor than personnel in "easier" assignments, such as suburban or alternative schools.
   b. *Supervisory responsibility.* The more extensive the supervisory responsibilities of a teacher or administrator, the greater the salary.
   c. *Academic degree.* Personnel with higher degrees will be rewarded accordingly by higher salary portions.
   d. *Work experience.* Salary should increase as the years of experience increase.
   e. *Additional college credits.* The more credits beyond a formal degree a district employee accumulates, the greater his salary.

   The above evaluation factors are not meant to be exhaustive. The criteria each district uses for salary schedules will be different.
3. Because the district has been plagued in the past by high teacher turnover, it has been decided to set up the salary schedule in order to maximize teachers' salaries. The top salary a teacher can earn in the district will be the maximum possible one under the current budget. In this way, the district hopes to provide incentive to teachers to remain within the district.

Your problem is to set up a salary schedule which conforms to all the above criteria, under the additional assumption that the current budget allows 18 million dollars for teachers' and administrators' salaries.

Mathematical Model

The measure of effectiveness for this problem will be the amount of money the highest paid teacher can earn. We want to maximize this amount in order to fulfill condition 3 above.

We must consider five evaluation factors in determining salaries: learning environment, supervisory responsibility, academic degree, work experience, and additional college credits. Personnel will be ranked within each factor according to their positions. There will be a minimum or basic amount of compensation associated with the lowest rank for each factor. As a person's rank increases within the factor, his compensation from that factor will increase according to the relative weight associated with his rank. The controllable variables will be the minimum compensations for each evaluation factor. There are five evaluation factors, so we will need five controllable variables, e, d, w, s, and c.

To illustrate how the compensations from factors will work, consider the factor of learning environment. In this district, teachers or administrators may be ranked as being in a difficult, medium, or easy learning environment, corresponding to whether they are associated with inner city, suburban, or alternative schools respectively. (These categories were agreed upon during negotiations.) If a person receives a rank of easy for learning environment (that is, works in an alternative school), he or she will receive a certain amount of compensation, which we shall indicate e. This amount, e, is one of our controllable variables. If the ranking is medium—that is, if he or she works in a suburban school—it has been established that he or she shall receive *twice* the minimum or basic compensation, or $2e$. A rank of difficult (that is, a position in inner-city schools) will entitle the person to receive *three times* the minimum amount, or $3e$. Therefore, the relative weights associated with the categories of difficult, medium, and easy learning environments are, respectively, 3, 2, and 1.

The amount of money a person receiv cording to the difficulty of the learning ment is only part of the total salary. T the salary will be determined by the ar or she is eligible for in each of the other four evaluation factors.

A summary of the evaluation factors, the notation for the minimum or basic compensation for each factor (our controllable variables), and the relative weights for determining salary within each factor is given in Table 3-10. For now, disregard the last two columns of numbers; their significance is discussed below. Remember that the relative weights associated with the various categories within each factor are assumed to have been agreed upon in negotiation. It should be pointed out also that whole-number weights are not necessary. For example, if it had been specified that teaching in a difficult learning environment would pay three and one-half times the basic amount for that factor, or $3.5e$, the relative weight for this category would be 3.5. We have chosen whole-number weights in this problem merely for simplification.

Now that we have defined the controllable variables and the measure of effectiveness, we must formulate the mathematical model. The measure of effectiveness is the highest possible salary a teacher can earn. Referring to Table 3-10, we see that a teacher having the following ranks within the five evaluation factors can earn the highest possible salary:

Evaluation Factor	Rank and Corresponding Weight	Salary from Each Factor
Learning environment	Difficult (3)	$3e$
Supervisory responsibility	Class-wide (2)	$2s$
Highest degree	Ph.D. or Ed.D. (5)	$5d$
Work experience	13+ years (7)	$7w$
Additional credits	29+ credits (5)	$5c$

The total salary (symbolized by T) that a top teacher can earn, then, would be:

$$3e + 2s + 5d + 7w + 5c = T$$

We want to find values for the controllable variables e, s, d, w, and c such that T is a maximum.

Now we must express the constraints on the controllable variables. First, let's assume that during salary negotiations certain guidelines were established for the minimum or basic amounts of compensation for each factor. For example, it was agreed that e, the basic amount paid because of learning environment, will be somewhere between \$200 and \$1000. (The lower limit was probably negotiated by the teachers, the upper limit imposed by the school board.) We need two mathematical expressions to indicate these limits on e:

$e \leq 1000$ (e will be at most \$1000)

$e \geq 200$ (e will be at least \$200)

Similarly, other guidelines were set up for s, d, w, and c. These and the corresponding mathematical equivalents are given in Table 3-11.

Other constraints on these controllable variables are that we have been given lower limits on salaries for each of the five district functions. Recall that these lower limits are:

District Function	Lowest Salary
Superintendent	\$22,000
Administrator	15,800
Department Head	11,500
Teacher	8,500
Teacher Aide	6,000

The district has definite criteria regarding the minimum qualifications for each district function. For example, the teacher who is paid the least amount would have the following characteristics within each evaluation factor and corresponding salary from each factor:

Evaluation Factor	Rank and Corresponding Weight	Salary from Each Factor
Learning environment	Easy (1)	$1e$
Supervisory responsibilities	Class-wide (2)	$2s$
Highest degree	B.A. (2)	$2d$
Work experience	0-2 years (1)	$1w$
Additional credits (1)	0-7 credits (1)	

Therefore, the least possible salary a teacher could be paid in this district is:

$$1e + 2s + 2d + 1w + 1c$$

Furthermore, we know this amount must be no smaller than \$8500. Therefore, one of the constraints on our controllable variable is:

$$1e + 2s + 2d + 1w + 1c \geq 8500$$

Similarly, we can formulate mathematical expressions for the minimum salaries for each of the other district functions. Table 3-12 tabulates the minimum requirements for each of

Evaluation Factor	Controllable Variable	Relative Weight and Categories	Number of Employees in This Category	Weighted Total
Learning environment	e	3 difficult	220	660
		2 medium	1166	2332
		1 easy	0	0
				2992e
Supervisory responsibility	s	7 single district wide (Superintendent)	1	7
		6 district wide (Assistants to superintendent)	5	30
		5 single school wide (Principals)	5	25
		4 school wide (Vice-principals, etc.)	25	100
		3 department wide	90	270
		2 class wide	1200	2400
		1 student wide (Teacher aides)	60	60
				2892s
Highest academic degree attained	d	5 Ph.D. or EdD	20	100
		4 M.A.	120	
		3 M.Ed.	1100	
		2 B.A. or B.S.	1100	
		1 A.A.	4[illegible]	
				6126d
Work experience	w	7 13 + years	16	112
		6 11 - 12 years	100	600
		5 9 - 10 years	300	1500
		4 7 - 8 years	400	1600
		3 5 - 6 years	500	1500
		2 3 - 4 years	50	100
		1 0 - 2 years	20	20
				543[illegible]
Additional college credits	c	5 29 + credits	600	3000
		4 22 - 28 credits	500	2000
		3 15 - 21 credits	200	600
		2 8 - 14 credits	50	100
		1 0 - 7 credits	36	36
				5736c

Table 3-10.
Evaluation factors in salary schedule model.

Evaluation Factor	Controllable Variable	Limits	Mathematical Statements
Learning environment	e	At most $1000 At least $200	$e \leq 1000$ $e \geq 200$
Supervisory responsibility	s	At most $3000 At least $500	$s \leq 3000$ $s \geq 500$
Highest degree	d	At most $2000 At least $200	$d \leq 2000$ $d \geq 200$
Work experience	w	At most $2000 At least $100	$w \leq 2000$ $w \geq 100$
Additional credits	c	At most $500 At least $100	$c \leq 500$ $c \geq 100$

Table 3-11.
Guidelines for evaluation factor values.

these functions; the last row indicates the minimum salary for each and its corresponding lower limit. The maximum possible salaries have also been formulated, although they will not be part of the constraints since we have not been given any upper limits on salaries.[12]

Summarizing from the information in Table 3-12, we find the constraints on minimum salaries for the five district functions to be as follows:

[12] It is difficult to set a maximum salary for a teacher aide. We have assumed here that the highest paid teacher aide would work in a difficult learning environment and would have student-wide responsibilities, no more than a B.A. or B.S. degree, 13+ years of experience, and no more than 7 credits beyond the Bachelor's degree.

Evaluation Factor		DISTRICT FUNCTIONS: Superintendent	Administrator	Dept. Head	Teacher	Teacher Aide
Learning environment (e)	Max	Difficult [illegible]	Difficult [illegible]	Difficult (3)	Difficult [illegible]	Difficult (3)
	Min	Easy (1)	Easy (1)	Easy (1)	[illegible]	Easy (1)
Supervisory responsibility (s)	Max	[illegible] district [illegible]	Single school [illegible]	Department (3)	[illegible] (2)	Student (1)
	Min	[illegible]	School (4)	Department (1)	Class (2)	Student (1)
Highest academic degree (d)	Max	[illegible]	Ph.D., EdD [illegible]	[illegible], EdD [illegible]	Ph.D., EdD [illegible]	B.A. [illegible]
	Min	M [illegible]	M.Ed. [illegible]	B.A. [illegible]	B.A. [illegible]	A.A. [illegible]
Work experience (w)	Max	[illegible] years [illegible]	[illegible] years [illegible]	[illegible] years (2)	[illegible] years	[illegible] years (2)
	Min	[illegible]	[illegible]	[illegible] years (1)	[illegible] (1)	[illegible] years (1)
Additional college credits (c)	Max	[illegible] credits [illegible]	[illegible] credits (5)	[illegible] credits (5)	[illegible] credits	[illegible] 7 credits (1)
	Min	[illegible] credits (1)	[illegible] credits [illegible]	[illegible] credits (1)	[illegible] credits (2)	0 - 7 credits (1)
Maximum salary		[illegible] + 5w [illegible]	[illegible] + 5w [illegible]	3e + [illegible] + 5d + 2w + 5c	[illegible] + 5d + 2w + [illegible]	3e + 1s + 2d + 2w + 1c
Minimum salary		1e + [illegible] + w + [illegible]	1e + 4s + 3d + 1w + 1[illegible]	[illegible] + 1w + 1[illegible]	1e + [illegible] + 2d + 1w + 1[illegible]	1e + 1s + 1d + 1w + 1c
		[illegible]	15800	11500	8500	6000

Table 3-12.
Minimum requirements for five district functions.

District Function	*Mathematical Statement of Least Possible Salary*
Superintendent	$1e + 6s + 3d + 2w + 1c \geqslant 22{,}000$
Administrator	$1e + 4s + 3d + 1w + 1c \geqslant 15{,}800$
Department Head	$1e + 3s + 2d + 1w + 1c \geqslant 11{,}500$
Teacher	$1e + 2s + 2d + 1w + 1c \geqslant 8{,}500$
Teacher Aide	$1e + 1s + 1d + 1w + 1c \geqslant 6{,}000$

The only constraint left to form on the controllable variables is the total available budget for teachers and administrators' salaries—eighteen million dollars. Referring again to Table 3-10 notice that the fourth column tabulates the number of persons within the district who are in each category of the evaluation factors. For example, 220 persons work in "difficult" learning environments, 1166 persons are in "medium" difficulty learning environments, and no persons are in "easy" learning environments. The 220 people in the "difficult" environment will each have to be paid three times the basic amount for the learning environment factor, and the 1166 in "medium" environments twice the basic amount. So, in all, the amount of money necessary for salaries in the learning environment factor is:

$$(220 \cdot 3 + 1166 \cdot 2) \cdot e = (660 + 2332) \cdot e = 2992e$$

Table 3-10 calculates similar figures for each of the other four evaluation factors, with the following results:

Evaluation Factor	*Money Necessary for Salaries*
Learning environment	$2992e$
Supervisory responsibility	$2892s$
Highest academic degree	$6126d$
Work experience	$5432w$
Additional college credits	$5736c$

Since the budget is limited to 24 million dollars for salaries, the mathematical expression of this constraint on the total money necessary for salaries is:

$$2992e + 2892s + 6126d + 5432w + 5736c \leqslant 24000000$$

The entire mathematical model for this problem is summarized as follows:

$$\left.\begin{array}{l} e \leqslant 1000 \\ s \leqslant 3000 \\ d \leqslant 2000 \\ w \leqslant 2000 \\ c \leqslant 500 \end{array}\right\} \text{upper limits for values of minimum or basic amounts for each evaluation factor}$$

$$\left.\begin{array}{l} e \geqslant 200 \\ s \geqslant 500 \\ d \geqslant 200 \\ w \geqslant 100 \\ c \geqslant 100 \end{array}\right\} \text{lower limits for values of minimum or basic amounts for each evaluation factor}$$

(both groups: constraints)

	1st var (e)	2nd var (s)	3rd var (d)	4th var (w)	5th var (c)	Sign	Right-hand side
1st constraint	1	0	0	0	0	≤	1000
2nd constraint	0	1	0	0	0	≤	3000
3rd constraint	0	0	1	0	0	≤	2000
4th constraint	0	0	0	1	0	≤	2000
5th constraint	0	0	0	0	1	≤	500
6th constraint	2992	2892	6126	5432	5736	≤	24000000
7th constraint	1	0	0	0	0	≥	700
8th constraint	0	1	0	0	0	≥	500
9th constraint	0	0	1	0	0	≥	200
10th constraint	0	0	0	1	0	≥	100
11th constraint	0	0	0	0	1	≥	100
12th constraint	1	6	3	2	1	≥	22000
13th constraint	1	4	3	1	1	≥	15800
14th constraint	1	3	2	1	1	≥	11500
15th constraint	1	2	2	1	1	≥	8500
16th constraint	1	1	1	1	1	≥	6000
Object function	3	2	5	7	5		

Table 3-13.
Salary schedule problem prepared for input to LPRG.

$$\left.\begin{array}{l} e + 6s + 3d + 2w + c \geq 22000 \\ e + 4s + 3d + w + c \geq 15800 \\ e + 3s + 2d + w + c \geq 11500 \\ e + 2s + 2d + w + c \geq 8500 \\ e + s + d + w + c \geq 6000 \end{array}\right\} \text{minimum possible salaries for each district function}$$

$$\left.2992e + 2892s + 6126d + 5432w + 5736c \leq 24000000\right\} \text{budget}$$

(all of the above: constraints)

Find values for e, s, d, w, and c subject to the above constraints and such that the top teacher's salary is a maximum:

$$3e + 2s + 5d + 7w + 5c = T \} \quad \text{object function}$$

We are now ready to assign an order to the variables and the constraints and to prepare the model for input to LPRG. The variables have always been referenced in the order e, s, d, w, c, so we will agree to use that order. As usual, we must have all the "less-than" constraints first, the "equality" constraints second (there are none), and the "greater-than" constraints third. We will agree that the upper limit constraints are the first five, the budget constraint is the sixth, the lower limit constraints are the seventh through eleventh, and the minimum salary constraints are the twelfth through sixteenth. As the upper and lower limit constraints have only one of the variables in them, they will have to be rewritten using 0 coefficients for the missing variables. Table 3-13 summarizes the salary schedule problem as we have prepared it for input to LPRG.

Be careful when you input data involving large numbers. Use commas only to separate numbers: never use them within a number. For example, you would *not* type the number 2992 as "2,992"; the correct coding would be "2992."

Analysis of the Solution

The computer listing for the solution to the salary schedule problem is given in Figure 3-13 with the following results (rounded to the nearest penny):

Evaluation Factor	*Controllable Variable*	*Value*
Learning environment	e	524.59
Supervisory responsibility	s	3000
Highest academic degree	d	400
Work experience	w	1975.41
Additional credits	c	100

The corresponding highest possible salary for a teacher, which we maximized, is $23,901.6.

Using the equations in the last row of Table 3-12 and the above values for the controllable variables, we can calculate the lowest and highest possible salaries for each district function:

```
IF MAXIMIZING THE OBJECT FUNCTION, TYPE 1.
IF MINIMIZING THE OBJECT FUNCTION, TYPE -1.
?1

NUMBER OF VARIABLES?   ?5

NUMBER OF CONSTRAINTS?   ?16

NUMBER OF LESS-THAN CONSTRAINTS?   ?6

NUMBER OF EQUALITY CONSTRAINTS?   ?0

NUMBER OF GREATER-THAN CONSTRAINTS?   ?10

ENTER THE COEFFICIENTS OF THE CONSTRAINTS, SEPARATED BY COMMAS.

CONSTRAINT  1    ?1,0,0,0,0

CONSTRAINT  2    ?0,1,0,0,0

CONSTRAINT  3    ?0,0,1,0,0

CONSTRAINT  4    ?0,0,0,1,0

CONSTRAINT  5    ?0,0,0,0,1

CONSTRAINT  6    ?2992,2892,6126,5432,5736

CONSTRAINT  7    ?1,0,0,0,0

CONSTRAINT  8    ?0,1,0,0,0

CONSTRAINT  9    ?0,0,1,0,0

CONSTRAINT  10   ?0,0,0,1,0

CONSTRAINT  11   ?0,0,0,0,1

CONSTRAINT  12   ?1,6,3,2,1

CONSTRAINT  13    1,4,3,1,1

CONSTRAINT  14   ?1,3,2,1,1

CONSTRAINT  15   ?1,2,2,1,1

CONSTRAINT  16   ?1,1,1,1,1
ENTER THE RIGHT-HAND SIDES OF ALL THE CONSTRAINTS, SEPARATED BY COMMAS.
?1000,3000,2000,2000,500,24000000,200,500,2[illegible]0,100,100,22000,15800,11500,8G

ENTER THE COEFFICIENTS OF THE OBJECT FUNCTION, SEPARATED BY COMMAS.
?3,2,5,7,5

DO YOU WISH TO CHANGE OR CORRECT ANY OF THE CONSTRAINTS?
(YES = 1, NO = 0)?0

DO YOU WISH TO CHANGE OR CORRECT THE RIGHT-HAND SIDES?
(YES = 1, NO = 0) ?0

DO YOU WISH TO CHANGE OR CORRECT THE OBJECT FUNCTION?
(YES = 1, NO = 0) ?0


ANSWERS:
VARIABLE        VALUE
 1                524.589
 2                3000.
 3                400
 4                1975.41
 5                100

THE VALUE OF THE OBJECT FUNCTION IS  25901.6
THIS VALUE IS A MAXIMUM.

DONE
```

Figure 3-13.
Salary Schedule Problem Solved Using LPRG.
In this case there is not enough space on one line to type all the right-hand sides and show each number one after another; however, LPRG is designed to continue to accept each number, even if the number appears on the paper to be typed directly over the preceding digit. Therefore, when the right-hand margin is reached, just continue typing as usual.

District Function	Lowest Possible Salary	Highest Possible Salary
Superintendent	23,775.41	28,9[illegible]1.64
Administrator	15,800.00	[illegible]901.64
Department Head	12,460.00	[illegible].901 64
Teacher	9,400.00	23,901.64
Teacher Aide	6,000.00	19,301.64

By substituting the values for the controllable variables in the original constraints, we can see that all the original constraints have been satisfied. For example, the total amount necessary for salaries is:

$$2992e + 2892s + 6126d + 5432w + 5736c$$
$$= 2992(524.59) + 2892(3000) + 6126(400) + 5432(1975.41) + 5736(100)$$
$$= 1569573.28 + 8676000 + 2450400 + 10730427.12 + 573600$$
$$= \$24000000.40$$

Notice that this figure is 40¢ over the allowable budget of $24 million. This error arose not because LPRG calculated incorrectly, but from rounding. Further, the excess is so small that it does not really represent a problem. Surely one person (perhaps the superintendent?) will agree to accept a dock in pay of [illegible]

One of the major adva[illegible] using this type of salary schedule i[illegible] completely objective: given any perso[illegible] position and background, we can quickly calculate the salary. For example, the salary of a principal (administrator, single school responsibility) in a suburban school with an M.Ed., 7 years' experience, and 16 credits beyond his or her degree would be:

$$2e + 5s + 3d + 4w + 3c$$
$$= 2(524.59) + 5(3000) + 3(400) + 4(1975.41) + 3(100)$$
$$= 1049.18 + 15000 + 12000 + 7901.64 + 300$$
$$= \$25[illegible]50.82$$

Refer to Table 3-10 for the relative weights of the controllable variables used above.

Although we will not go into detail here, it is possible to introduce many more constraints into the salary schedule problem. Consider the following list of possible modifications.

1. The salary schedule could be based on as many evaluation factors as your computer program can handle. Such criteria as quality of teaching, inservice training, additional workloads (for example, coaching), and priority of subject matter could all be included.
2. There are many possibilities for measures of effectiveness in this type of problem: for example,
   - Beginning teacher's salary (maximize);
   - Total budget (minimize);
   - Compensation for teaching in a difficult learning environment (maximize the quantity e in the previous problem);
   - Total amount allotted for work experience (maximize).
3. The model could be constructed to allow for certain salary spreads within each district function. For example, a constraint could be added specifying that the least possible salary for some district function was not less than a certain percentage of the highest possible salary for the same function. To be more specific, we could make the condition that the least possible teacher's salary was at least 50 percent of the highest possible teacher's salary. If this were $16,000, the lowest possible teacher's salary would have to be $8,000 or more in order to conform to this constraint.
4. You could allow for percentage salary overlaps between district functions. For example, it would be possible to specify that the highest paid teacher could earn at least 10 percent more than the lowest paid department head.
5. You could specify dollar spread between the highest salaries for district functions—for example, the highest-paid teacher must earn at least $4,000 more than the highest-paid teacher aide.
6. You could set upper limits on maximum salaries. For example, the maximum salary for teacher aides in the previous problem, $19,301.64, turned out to be fairly high. This figure is probably not realistic, and you might want to specify

a later computer run that the maximum teacher aide salary is no higher than, for example, $12,000.

A mathematical model for salaries as presented here is ideal for conducting simulation studies. An administrator could outline salary schedules for future years based on predictions of how the school budget will change. Or, he could assess future budgetary requirements based on the need for more teachers within the district. The possibilities are endless.

There is one major limitation in this salary schedule model. With the large number of constraints on only a few controllable variables, it is very easy to unknowingly construct a problem which has no solution at all, much less a "best" solution. Such a situation would have arisen in the original salary schedule problem if we had had a total budget of only 4 million dollars. With the constraints we put on minimum values for e, s, d, w, and c it would not be possible to keep within this budget. We can easily see this by substituting the minimum values for the controllable variables in the budget constraint. This results in a figure greater than 4 million:

$$
\begin{aligned}
2992e + 2892s + 6126d + 5432w + 5736c &= 2992(200) + 2892(500) + 6126(200) \\
&\quad + 5432(100 + 5736(100) \\
&= 598400 + 1446000 + 1225200 \\
&\quad + 543200 + 573600 \\
&= 4386400
\end{aligned}
$$

The computer program LPRG will tell you when there are no solutions to a linear programming problem. If that happens, you can vary some of the constraints until there *is* a solution. The process of adjusting some constraints in order to find a solution should in itself provide valuable information; an administrator may find, for example, that demands from the teachers' union are impossible to meet under the current budget.

REVIEW

Exercises

1. (a) Re-solve the salary schedule problem assuming that your budget is cut by 10 percent. That is, you have been allotted $21.6 million for salaries instead of $24 million. Do you expect that the resulting figures in the salary schedule will all be reduced by 10 percent when compared to the original salary schedule?
   (b) Re-solve the problem assuming your budget has been increased to $27 million. What happens to the salary ranges for the district functions? Where does the extra money seem to be channeled?
2. Suppose you have a budget of 27 million dollars, but you want to produce a salary schedule in which the salaries of teachers and administrators are not as divergent as in problem 1 (b) above. What categories in which evaluation factors could you reweight in order to accomplish this revision, and how would you reweight them? What other constraints could be included or changed? Try the problem on the computer after you have made your "predictions." (There is no set answer to this question.)

THE MATHEMATICS BEHIND LINEAR PROGRAMMING

Much of mathematics makes use of the basic concepts of equations and inequalities to describe relationships between quantities (variables). Mathematical modeling is the practice of defining such relationships precisely by using equations and inequalities. Once this has been accomplished, all the power and sophistication of present-day mathematics can be brought to bear on the model, and solutions can be found. Linear programming is an appropriate problem-solving technique when all the equations and/or inequalities in the accompanying mathematical model are of a specific form—namely, linear—and when there is a quantity to be optimized.

LINEAR EQUATIONS AND THEIR GRAPHS

A linear equation or inequality is one in which the variables are added and/or subtracted from each other or from numbers; variables may not be multiplied or divided by each other or by themselves. Numbers (constants) may be added, subtracted, multiplied, and/or divided by variables or by themselves. The general form of a linear equation is:

$$Aa + Bb + Cc + \ldots + Xx + Yy + Zz = 0$$

where the upper-case letters represent constants and the lower-case letters are variables. Linear inequalities have the same form except that the equality sign is replaced with an inequality sign.

Linear equations get their name from the fact that graphs of their solutions are lines.[13] Let's look at an example. Consider the equation $2a + b = 0$. A graph of this equation is the set of all pairs of numbers which when substituted to a and b make the equation true. If we let $a = 1$ and $b = -2$ and substitute these values in the equation, we get

$$2a + b = 2(1) + (-2) = 2 - 2 = 0$$

which is a true statement. If we let $a = 1$ and $b = 1$ and substitute these values, we get

$$2a + b = 2(1) + (1) = 2 + 1 = 0$$

which is not a true statement. Therefore, we say that $a = 1$ and $b = -2$ is a solution to the equation and $a = 1$ and $b = 1$ is not a solution. There are many other solutions to this equation; in fact, there are an infinite number of them. Some are listed below.

Value of a	Value of b
3	-6
2	-4
1	-2
0	0
-1	2
-2	4
-3	6

In order to graph this equation, we must have a way to draw pairs of numbers. The standard method is to use a grid with a central point where each point on the grid represents one pair of numbers. The central point is called the origin of the grid and represents the pair of numbers (0,0). Any other point on the grid is defined in relationship to the origin. The first number of the pair (called the first or horizontal coordinate) tells how far the point is from the origin in the horizontal direction. If the first coordinate is positive, the point is to the right of the origin; negative, it is to the left. The second number (called the second or vertical coordinate) tells how far the point is from the origin in the vertical direction. If it is positive, the point is above the origin; if negative, below. This graphing system is called a Cartesian coordinate system, named after its developer, René Descartes, a French philosopher and mathematician. Figure 3-14 is a graph of several points, using the Cartesian coordinate system.

Now we are in a position to begin graphing the equation $2a + b = 0$. First, let's graph the pairs of numbers in the list on page 00. We will use a for the horizontal coordinate and b for the vertical coordinate.

[13] Frequently in mathematics we talk about graphing equations and inequalities. Technically, we are graphing all the solutions to an equation (or inequality) rather than the equations themselves; but we will use these two expressions interchangeably.

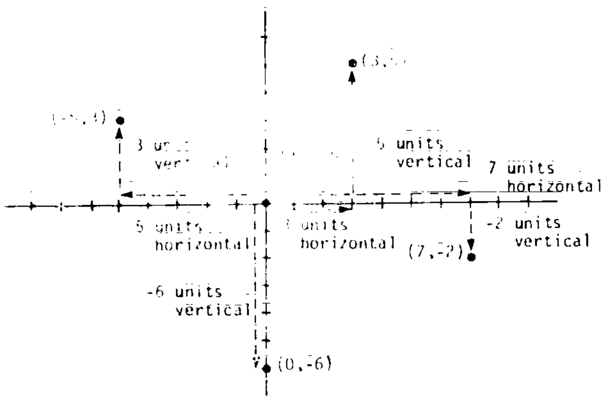


Figure 3-14.
Random points on a Cartesian Coordinate System.

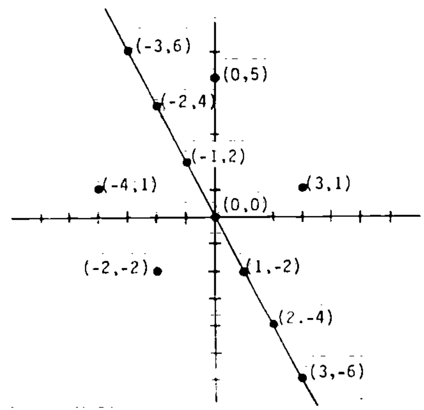


Figure 3-15.
Graph of Solution to $2a + b = 0$.

Notice that all of the pairs lie on the same line. In fact, all other pairs which are solutions to this equation are on the same line and all points on the line represent pairs which are solutions to the equation.

REVIEW

Terminology

1. LINEAR EQUATION: an equation in which the variables are only added or subtracted from each other or from constraints.
2. SOLUTION OF LINEAR EQUATION: a pair of numbers which result in a true statement when substituted for the variables in the equation.
3. CARTESIAN COORDINATE SYSTEM: a system of naming points in a plane with pairs of numbers.
4. ORIGIN: the central point of a grid, representing the pair (0,0).
5. HORIZONTAL COORDINATE: the first number in a pair, denoting horizontal distance from the origin.
6. VERTICAL COORDINATE: the second number in a pair, denoting the vertical distance from the origin.

1. Find five pairs of numbers which are solutions to the equation $a + 2b + 1 = 0$. Graph these pairs on a grid (use a for the horizontal coordinate and b for the vertical coordinate) and then graph all the solutions to the equation.
2. Graph the equations $a + b = 0$ and $3a + b + 2 = 0$. Hint: as you may or may not remember from geometry, two points completely determine a line. Therefore, it is only necessary to find two points on a line (that is, two solutions to an equation) in order to draw or graph that line. An easy way to find two solutions is as follows. Arbitrarily set one variable to zero (say $a = 0$) and find out the value of the second variable. Then set the second variable to zero ($b = 0$) and find the value of the first. You now have two solutions to the equation.

GRAPHING INEQUALITIES

Graphing inequalities is an easy matter once you know how to graph equations. Consider the inequality $2a + b \leq 0$. We are looking for all of the points (pairs of numbers) which make this statement true when the values for a and b are substituted into the inequality. Certainly all of the pairs of numbers on the line $2a + b = 0$ also satisfy this inequality; but there are many others. Let's pick out a point which is not on the line, say $(-4,1)$, and see if it is a solution.

$$2a + b = 2(-4) + (1) = -8 + 1 = -7 \leq 0$$

The result is a true statement so $(-4, 1)$ is a solution. Let's try $(-2, -2)$:

$$2a + b = 2(-2) + (-2) = (-4 + (-2) = -6 \leq 0$$

The result here is also a true statement. Let's try another point, say $(3,1)$.

$$2a + b = 2(3) + (1) = 6 + 1 = 7 \leq 0$$

The result is a false statement, so this point is not a solution to the inequality. Let's try one last point, $(0,5)$.

$$2a + b = 2(0) + (5) = 0 + 5 = 5 \leq 0$$

This also results in a false statement and so $(0,5)$ is not a solution. Let's look at our four trial points on a graph. Notice that both of the points off the line which are solutions to the inequality lie in the half-plane which is below the line, and that both of the points off the line which are not solutions to the inequality lie in the half-plane which is above the line. If more extensive testing were conducted, we would find that every point in the half-plane below the line is a solution to the inequality. Therefore we can graph the inequality in the following manner.

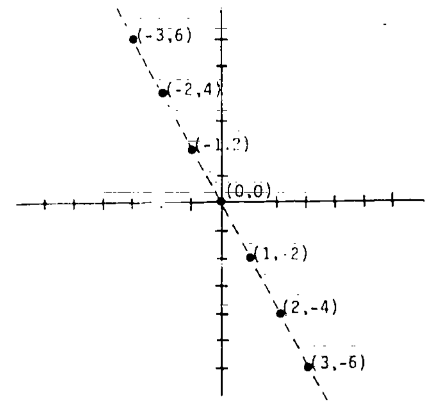


Figure 3-16.
Graph of $2a + b = 0$ plus Selected Points.

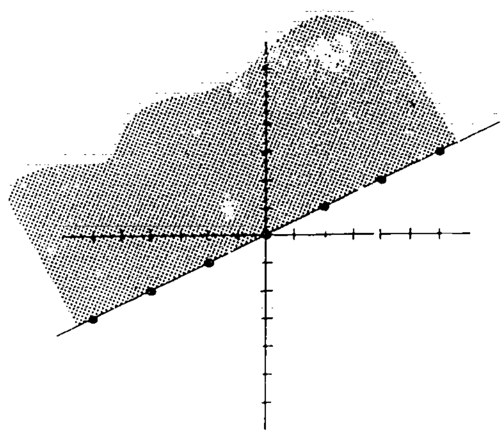


Figure 3-17.
Graph of $2a + b \leq 0$.

REVIEW

Exercise

1. Graph the following inequalities:
   a. $2b \leq 0$
   b. $3a + 1 \leq 0$
   c. $a + b \leq 0$
   d. $a - 4b$ [illegible] $1 \leq 0$

Hint: once the line is graphed, it is necessary to test one point on either side of the line. Whichever point is a solution lies on the side which should be shaded.

SOLUTION OF LINEAR PROGRAMMING PROBLEMS BY GRAPHING

Now that we have reviewed basic techniques for graphing linear equations and inequalities, we are ready to turn our attention toward solving linear programming problems. Let's return to our original TV set problem. Recall that the constraints were:

$$p \leqslant 60$$
$$c \leqslant 40$$
$$2p + 3c \leqslant 150$$

and the object function to be maximized was:

$$20p + 25c = T$$

Recall that we assumed the number of both consoles and portables made would be positive, since it makes no sense to produce −10 consoles. For the purposes of our discussion here, we shall have to make the non-negativity conditions explicit in the form of two extra constraints:

$$p \geqslant 0$$
$$c \geqslant 0$$

We are looking for pairs of numbers which satisfy all the constraints (the sc[illegible] to the problem). Then, from all the solu[illegible], we wish to pick the one(s) which gives a m[illegible] value to the object function. It will b[illegible] find all the solutions using graphing techniques. We do it by graphing all the constraints and finding the points which are on all of the graphs. Since they are on all of the graphs, they are solutions to each of the inequalities and therefore are solutions to the problem, since they satisfy each of the constraints. In graphing the constraints, we will use p as the horizontal coordinate and c as the vertical coordinate.

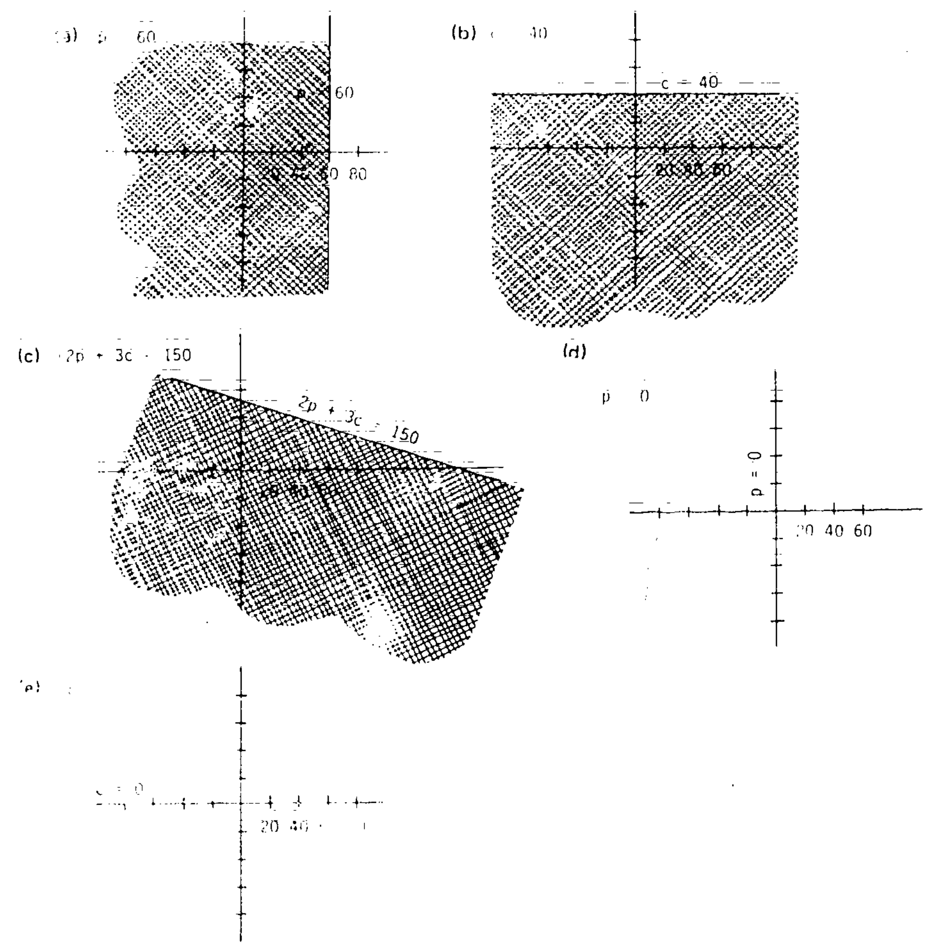


Figure 3-18.
Constraints of Television Production Problem.

If we superim... the five graphs onto one gr... and indicate the poi... which are on all five, ... have:

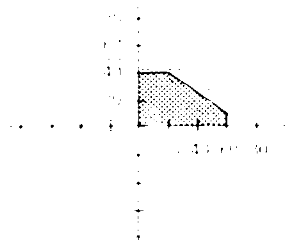


Figure 3-19.
Graph of Solutions to Television Production Problem.

Now we turn our attention toward finding the optimal solution. Suppose that we arbitrarily set the value of the object function at 2000 and graph the equation $20p + 25c = 2000$.

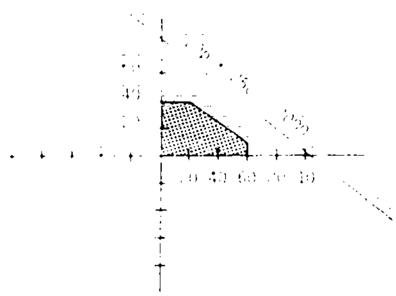


Figure 3-20.
Graph of Solutions with Object Function Value of 2000.

Notice that none of the solutions to the problem lie on this line. We can conclude from this that no pair of numbers which satisfy all the constraints will yield a value of 2000 for the object function.

Let's set the object function value at 1000 and examine the graph in Figure 3-21.

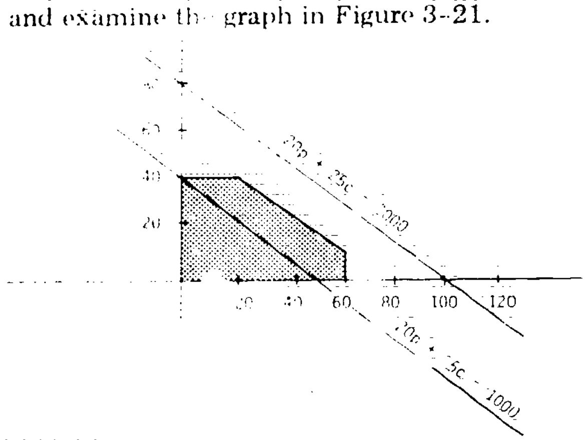


Figure 3-21.
Graph of Solutions and New Object Function Values of 1000.

Notice that there are many solutions which yield a profit of \$1000. Notice also the relative positions of the two lines representing different values of the object function. The lines are parallel to each other, and the line with a value of 1000 is under the line with a value of 2000. Where do you suppose the graph of the line $20p + 25c = 1500$ will fall? Let's graph it and see.

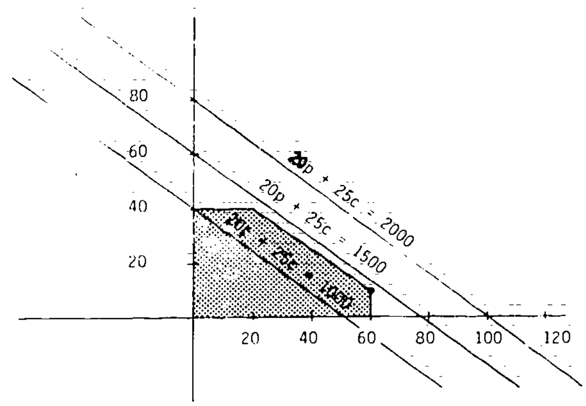


Figure 3-22.
Graph of Solutions and New Object Function Value of 1500.

Again, there are no solutions which will yield a profit of \$1500. Also, notice that this third line is parallel to the other two and lies between them. By this time you should begin to suspect the fact th... every new value for the object function will produce a new line which is parallel to the others. Graphically, then, we are looking for the highest line which is parallel to the three we have already generated and which intersects the graph of the solutions in at least one point. It should be clear from the graph, that the desired line will go through the upper right-hand corner, which is the point (60,10). This is the optimal solution, as we already know, and the equation of the line we are looking for is $20p + 25c = 1450$.

The above discussion was presented so that the user of these materials has a chance to develop an intuitive notion of what linear programming problems are all about. It is not intended to be a complete discourse on the mathematical technique of linear programming. There are several other methods of solving linear programming problems. The program LPRG uses one of them (called the two-phase simplex method) rather than the graphing method used above. But all of the methods accomplish the same thing, namely the identification of all possible

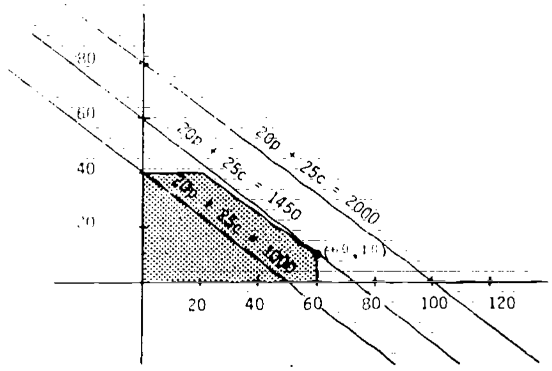


Figure 3-23.
Graph Showing Optimal Solution.

solutions to the problem and the subsequent identification of the optimal solution.

OVERVIEW OF LINEAR PROGRAMMING

ADVANTAGES AND DISADVANTAGES

We have pointed out several disadvantages and limitations to the decision strategy of linear programming, along with obvious advantages for the educational administrator. This section summarizes these.

Advantages

The primary advantage of using any operations research technique such as linear programming is that it provides *a rational basis for decision making*. Any opinion or decision is more believable and more valid if it can be substantiated by hard data. Linear programming provides definite criteria for determining the "best" solution to a problem.

Other advantages are:

1. Linear programming provides *a format for systematic analysis* of problems which are concerned with allocating resources in order to optimize a certain result. It forces the administrator to identify those variables over which he has control and the accompanying restrictions on those variables, and to analyze the interrelationships among them.
2. Using linear programming encourages *the identification of goals* and objective expression of these goals. The administrator is forced to establish the extent to which each variable affects the goal. Quantification of goals is invaluable in terms of evaluation and accountability, since it helps establish whether goals have been attained and how well they have been attained. Such questions as whether educational goals have been met are of vital importance to the administrator, since he is answerable to many groups, including the federal government, the state government and of course the taxpayers.
3. Once the problem has been expressed mathematically, linear programming computer programs such as LPRG provide the means for obtaining *fast answers at little cost*.
4. Programs like LPRG also provide a *means for simulating changes* in the linear programming problem in order to see how the goal and the controllable variables would be affected if such a change were implemented.

Disadvantages

The foremost disadvantage to using linear programming is the *difficulty in the mathematical formulation* of real-world situations. Some problems just do not lend themselves to mathematical expression. With others, successful mathematical modeling is limited by available mathematical techniques, variables that are unknown, and goals that are difficult to express quantitatively. Even when a problem can be expressed mathematically, the assumptions which were necessary for the formulation may have so changed the problem that the results are not useful. As computer programmers are prone to say, "Garbage in, garbage out." That is, if the data which go into a linear programming problem are not valid or representative of the real-world situation, the output from the problem will neither be correct nor useful.

Linear programming only *aids* in decision-making; it *does not actually make decisions.* Linear programming gives validation and support to decisions, but it cannot replace the human administrator who is ultimately responsible for determining courses of action.

SUMMARY OF THE MAIN POINTS IN LINEAR PROGRAMMING

Formulating Models

A. Recognition of the problem: Linear programming is an operations research tool used in problems where resources are to be assigned or allocated in order to optimize some result.
   1. Typical resources to be allocated in education problems:
      a. Money;
      b. Teachers and teacher aides;
      c. Facilities;
      d. Class time;
      e. Equipment.
   2. Typical results to optimize in education problems:
      a. Money: minimize costs; maximize certain salaries;
      b. Transportation time: minimize;
      c. Racial balance: maximize;
      d. Quality of instruction: maximize;
      e. Facilities: minimize empty classrooms and empty school bus seats.

B. Formulation of the mathematical model.
   1. Identification of the controllable variables (resources);
   2. Choice of a measure of effectiveness (quantity we wish to maximize or minimize);
   3. Mathematical representation of the object function (relate the controllable variables to the measure of effectiveness);
   4. Identification of the constraints (restrictions or limits on the values of the controllable variables);
   5. Mathematical representation of the constraints;
   6. Determination of the linearity of the model and hence the model's suitability for use of the linear programming technique.

Solving Linear Programming Problems

A. Types of solutions.
   1. Possible solution: satisfies all constraints;
   2. Optimal solution: satisfies all constraints and gives optimal (maximum or minimum) value to the object function.

B. Possibilities for solutions to linear programming problems.
   1. Many possible solutions but only one optimal solution;
   2. No solutions at all;
   3. An infinite number of optimal solutions which satisfy all constraints and give the same optimal result.

C. Solving linear programming problems using the computer program LPRG.
   1. Formulate mathematical model: listing "less-than" constraints, followed by "equality" constraints, followed by "greater-than" constraints, and object function;
   2. Prepare other necessary information for input to LPRG: whether object function is to be maximized or minimized; number of constraints; number of "less-than" constraints; number of "equality" constraints; number of "greater-than" constraints.
   3. Run the computer program.
   4. Analyze the results.

BIBLIOGRAPHY AND SUGGESTED REFERENCES

Anderson, D. R., D. J. Sweeney, and T. A. Williams. *Linear Programming for Decision Making: An Applications Approach.* St. Paul: Western Publishing Company, 1974.

Berrie, Phillip J. "A Linear Programming Model to Optimize District-Wide Assignment of Teachers." Ph.D. dissertation, University of Iowa, 1972.

Bruno, James E. "Alternative to the Use of Simplistic Formulas for Determining State Resource Allocation in School Finance Programs." *American Educational Research Journal, 6* (1969), 479-500.

——. "An Alternative to the Fixed Step Salary Schedule." *Education Administration Quarterly, 6* (1970), 26-46.

Decker, Michael, Issue Editor. "Education and Urban Society." Vol. III, No. 1, November, 1970.

Gass, Saul I. *An Illustrated Guide to Linear Programming.* New York: McGraw-Hill, 1970.

Harris, Roy D., Michael J. Maggard, and William G. Lesso. *Computer Models in Operations Research.* New York: Harper and Row, 1974. Suggested reading: Chapter 1, LINPRO, pp. 7-23.

Kwak, N. K. *Mathematical Programming with Business Applications.* New York: McGraw-Hill, 1973.

Luenberger, D. G. *Introduction to Linear and Nonlinear Programming.* Reading, Mass.: Addison-Wesley, 1973.

Matzke, Orville R. "A Linear Programming Model to Optimize Various Objective Functions of a Foundation Type State Aid Program." Ph.D. dissertation, University of Iowa, 1971.

McNamara, James F. "Mathematical Programming Models in Educational Planning," *Review of Educational Research, 41* (1971), 419-446.

——. "Operations Research and Educational Planning." *Journal of Educational Data Processing, 9* (1972), 1-12.

Mihram, G. Arthur. *Simulation, Statistical Foundations and Methodology.* New York: Academic Press, 1972.

Newton, Rita M., and Warren H. Thomas. "Design of School Bus Routes by Computer," *Socio-Economic Planning Science, 3* (1969), 75-85.

Ontjes, Robert L. "A Linear Programming Model for Assigning Students to Attendance Centers." Ph.D. dissertation, University of Iowa, 1971.

Reisman, Arnold, and Martin I. Taft. "A Systems Approach to the Evaluation and Budgeting of Educational Programs," *Socio-Economic Planning Science, 3* (1969), 245-277.

Schnittjes, Carl J. "Prediction of Success Through Linear Programming." Ph.D. dissertation, University of Iowa, 1971.

Shamblin, James E., and G. T. Stevens, Jr. *Operations Research, A Fundamental Approach.* New York: McGraw-Hill, 1974. Suggested reading: Chapter 10, Linear Programming, pp. 243-303.

Siemens, Nicolai, C. H. Marting, and Frank Greenwood. *Operations Research.* New York: The Free Press, 1973. Suggested reading: Chapter 5, Optimization Models for Planning Decisions, pp. 104-152.

Simmons, D. M. *Linear Programming for Operations Research.* San Francisco: Holden-Day, 1972.

Strum, Jay E. *Introduction to Linear Programming.* San Francisco: Holden-Day, 1972.

Vajda, S. *Theory of Linear and Nonlinear Programming.* London: Longman, 1974.

Van Dusseldorp, Ralph A., D. E. Richardson, and W. J. Foley. *Educational Decision-Making Through Operations Research.* Boston: Allyn and Bacon, 1971.

Wagner, Harvey M. *Principles of Operations Research.* Englewood Cliffs, N.J.: Prentice-Hall, 1975. Suggested reading: Chapters 2-6, pp. 33-222.

4 QUEUEING THEORY

CHAPTER PREVIEW

Many situations in education require waiting lines or a waiting period—for example, students waiting in line to register or equipment waiting for repair. To maximize efficiency in these situations, administrators often need to answer such questions as, Are the lines too long or waiting periods excessive? Are more or fewer facilities needed to correct problems or to maximize efficiency? Queueing theory provides a powerful tool for analyzing such situations and answering these questions.

This chapter presents the basic elements of queueing theory and illustrates its many uses in education. Emphasis is on analyzing realistic situations and interpreting statistics to provide the basis for effective decision making. An understanding of simple algebra will equip the reader to follow the step-by-step discussions of all hand calculations of queueing statistics. In addition, the use of computer programs to perform queueing analyses is explained and demonstrated. Exercises provide guided practice in using both hand calculations and computer programs.

CHAPTER AIMS

After completing this chapter, the reader will be able to use the terms specific to queueing theory and recognize which situations satisfy the conditions required for applying the theory. He will also be able to perform queueing theory analyses using either hand calculations or computer programs and to interpret the results for use in effective decision making.

INTRODUCTION TO QUEUEING THEORY

THE ELEMENTS OF QUEUEING THEORY

In everyday life there are many situations in which individuals must wait some length of time for a needed service. People commonly have to wait in line to buy a theatre ticket, to be seated in a restaurant, and to park their cars in a parking lot; airplanes regularly line up when arriving and departing; lines of cars at a turnpike toll booth are a common sight. Each of these examples involves a queueing or waiting system. Many other common situations involve similar problems: situations in which things must be placed in order (filing, stocking goods), for example, or in which a certain percentage of items must be repaired periodically so that back-up items are required.

The operations research technique called queueing theory provides a means for analyzing these queueing systems and obtaining statistics that describe their important features. In addition, the theory can be used to analyze the effects of making changes in the queueing system and determine whether such changes might be acceptable or beneficial. In effect, a decision maker can simulate changes before they are implemented. A bank manager might use queueing theory to tell him how much customer waiting time would be affected if he were to add another teller; the highway commission might use it to investigate the effects of constructing more toll booths at a turnpike interchange.

There are, of course, many situations in education that can be effectively dealt with using queueing theory. Students waiting to see a counselor, children waiting for appointments with a school district psychologist, schools or classrooms waiting for major repairs by a carpenter, and purchase orders waiting to be filled are all examples.

All waiting line situations have two characteristic parts:

1. A queue or waiting line of customers (e.g. teachers waiting to use projectors; students waiting to see a counselor).
2. Service facilities which provide a service or product to the customers (e.g. the projectors; the counselor).

In order to apply the techniques of queueing theory, some quantities must be known (or estimated) about the customers and the facilities. These include the number of potential customers, the number of service facilities available, the average number of customers that arrive in any given unit of time to use the facility, and the average length of time required to provide the service. For example, 300 high school students (customers) may be advised by one counselor (service facility). The students may wish to see the counselor at an average rate of about four students per day (arrival rate) and from the counselor's records it might be found that an average counseling session lasts 15 minutes (length of time required to provide service).

Once these quantities are known, queueing theory can provide informative and useful statistics: the average waiting time for an arriving customer, the average number of customers waiting in line, the probability that a given number of customers will be waiting for service at any given time, and the probability that the service facility is idle (i.e. that there are no customers being serviced or waiting in line). Thus queueing theory can help answer such questions as: How many people, on the average, are waiting to use the facility at any given time? How long should one expect to wait before the facility is available for use? Is the time people spend waiting so costly that another facility would be justified? Could scheduling or rescheduling the users eliminate any of the wasted idle time of the facility?

In addition to providing data to help answer questions like these, queueing theory gives the administrator a means for examining a problem from several different angles. Suppose purchase orders for a district seem to be continually stacked up and teachers are complaining about delays in obtaining materials. The orders are handled by one purchasing clerk. An administrator can use data from queueing theory to help decide whether the extra cost of a second clerk would be justified in terms of increased speed in processing purchase orders. Alternatively, the administrator might use queueing theory to determine how many clerks would be necessary to insure that most orders would be processed within a certain number of days. Or, the administrator might discover in the course of the analysis that the present clerk needs to be retrained or replaced. Or, again, the administrator might discover that far too many purchase orders are being filled and that it would be most efficient to anticipate the use of certain materials and have them on hand at the beginning of the school year.

It is obvious that queueing theory is a valuable decision strategy for the administrator concerned about efficient utilization of time and facilities. In summary, its four primary uses are the following:

It helps the administrator to study current queueing systems (or non-systems) and to pinpoint problem areas and potential solutions.

It provides statistics that can be used as criteria for determining the "best" solution to a problem. For example, the best solution to the purchase order problem discussed above might be the specific number of clerks (service facilities) required to insure that no purchase order must wait over two days to be processed.

It provides data which may suggest possible changes in the system being studied. Examination of the purchase order problem, for instance, may suggest such possible changes as varying the number of purchasing clerks (service facilities), reducing the frequency of purchase orders by requesting that teachers submit material requirements well before the school year begins, and reducing the amount of time required to process the orders by retraining the clerk.

It allows the administrators to test, at a small cost, several alternative modifications of the system and determine the most economical or practical course to pursue.

Queueing theory is a rapidly growing area of mathematical inquiry. Many types of models have been studied in great detail; this chapter will deal specifically with types of queueing situations that have a great many practical applications.

QUEUEING THEORY TERMS

Following are the basic terms used in queueing theory and their definitions.

Service Facility—location at which a service is rendered (purchasing clerk, machine to be used, repairman, school psychometrist). The letter F denotes the number of service facilities in a queueing situation.

Customer—user of the service facility. A customer need not be a person. A machine waiting for repair is considered a customer in queueing theory.

Source Field—the population of potential users or customers of the facility. Two types of source fields are considered in queueing theory:

A small, finite population, i.e., less than 30. The letter M denotes the exact size of such a population.

An infinitely large population. In general, large populations—greater than or equal to 30—will be close enough to an infinitely large population for use in practical problems.

Waiting Time—the time elapsed from when a customer joins the waiting line (queue) to the time when service is begun for that customer.

Arrival Rate—the average number of customers arriving at the service facility during a unit of time; denoted by A.

Service Rate—the average number of customers that can be serviced by one service facility during a unit of time, assuming no idle time; denoted by S.

To illustrate the use of these terms, let's consider the situation of a school district's purchasing clerk who receives purchase orders from all schools in the district. Typically, the time it takes to process orders varies because of differences in availability of items, suppliers' delivery schedules, and so forth.

Suppose the purchasing clerk keeps records for two weeks (ten working days) and obtains the following data:

Day	*No. of Orders Received*	*Hours Worked Each Day on Processing Orders*
1	1	2.0
2	20	8.0
3	6	5.7
4	7	6.5
5	8	8.0
6	3	5.0
7	19	7.5
8	9	8.0
9	13	8.0
10	14	8.0
	100 Orders	66.7 Hours

In this situation, the *customers* are the purchase orders. The *service facility* is the one clerk (F = 1). The *source field* is infinite, since a school district would presumably produce more than 30 purchase orders. Since the clerk received 100 orders in 10 working days, the *arrival rate* of customers is an average of 10 orders/day (A = 10 orders/day). The clerk worked 66.7 hours to process 100 orders, so we could say he worked at the *service rate* of 1.5 orders/hour

(S = 1.5 orders/hour). Since queueing theory requires that arrival and service rates be specified in the same time units, we will express the service rate in terms of orders per day. Assuming that the clerk works 8 hours a day,

1.5 orders/hr. × 8 hrs./day = 12 orders/day.

Therefore, S = 12 orders/day. The average waiting time of an order, which will not be calculated at this point, would be the average time elapsed from when an order arrived in the purchasing clerk's office to the time when the clerk began to process the order.

REVIEW

Terminology

1. QUEUE: waiting line of customers.
2. SERVICE FACILITY (F): location at which service is rendered.
3. CUSTOMER: user of service facility.
4. SOURCE FIELD: population of potential users; finite M is less than 30 (<30) and infinite is more than 30 (>30).
5. WAITING TIME: time from arrival at F until service begins.
6. ARRIVAL RATE (A): average number of arrivals during a unit of time.
7. SERVICE RATE (S): average number of customers served by one F during a unit of time.

Exercise

1. Assume that a school has one ditto machine for teacher use and that several teachers have complained that they frequently have to wait to use the machine because it is busy. In this situation, suppose the principal keeps track of the uses for 30 days and discovers that the machine was used by teachers 480 times and that the total time the machine was actually being used was 160 hours. For this situation, what is F (number of facilities)? What is A (arrival rate)? What is S (service rate)? Express both A and S in hours.

CONDITIONS FOR QUEUEING THEORY

In the queueing situation, certain conditions must be satisfied in order for the queueing theory described here to give useful results. These conditions arise because of the mathematics underlying the theory. They are described below.

1. *Arrivals of customers must be independent and occur at random time intervals.* Independence of arrivals means that an arrival is not affected by (is independent of) other arrivals. For example, if customers frequently came to use a facility but saw it was busy or that a long line was waiting and hence they left, those arrivals (or nonarrivals) would depend on other arrivals—the customers already using or waiting to use the facility—and the condition of independence would be violated. The fact that customers arrive at random time intervals means that arrivals do not follow a particular pattern. If arrivals were consistently most frequent at the beginning of the day or immediately after the change of class periods, or followed a regular schedule such as one event per half hour, this condition would be violated.
2. *Service times are assumed to be independent over time.* Independence of service times over time means that the service time for any one customer is assumed to be independent of the service time for any other customer and to be unaffected by such factors as the length of the line or the tiredness of workers; in other words, the service time is assumed neither to speed up because the line is longer nor to slow down because workers are tired.

3. *The queueing system is in equilibirum.* This means the facility has been operating long enough for the laws of probability on which queueing theory is based to take effect. In general, queueing theory does not apply when the facility first opens, usually at the beginning of a day. In some situations at the beginning of the day there will be no arrivals before the facility opens and the first arrival will always be served immediately. In other situations, anxious customers will line up before a facility opens and all will spend some time waiting to use the facility. Therefore, queueing theory does not apply when a facility is first opened, because the expected waiting time and number waiting will be different than after the facility has been open for a while.

Figure 4-1 illustrates a queueing system in equilibrium which involves cars in line at a service station. The drawing suggests that the length of the waiting line is fairly constant and customers arrive steadily. The amount of time car C has to wait in line until it begins service will be about the same time that car F or any other car must wait.

A simple example of a system which is not in equilibrium is a service station when it first opens. The first customer would encounter no waiting line—he would simply drive up to the gas pump and be serviced. His waiting time would be zero. This situation, however, applies only to the first few customers. After the station has been open a while, it is assumed that the arrivals will settle down to a constant rate; a waiting line will form which remains fairly stable; and the length of time each customer will have to wait for service will be quite similar. The system would then have reached equilibrium, and queueing theory could be applied.

4. *Customers are served on a "first come, first served" basis.* This condition requires that there are no priority customers who go to the head of the waiting line.

5. *Once a customer joins the queue, he does*

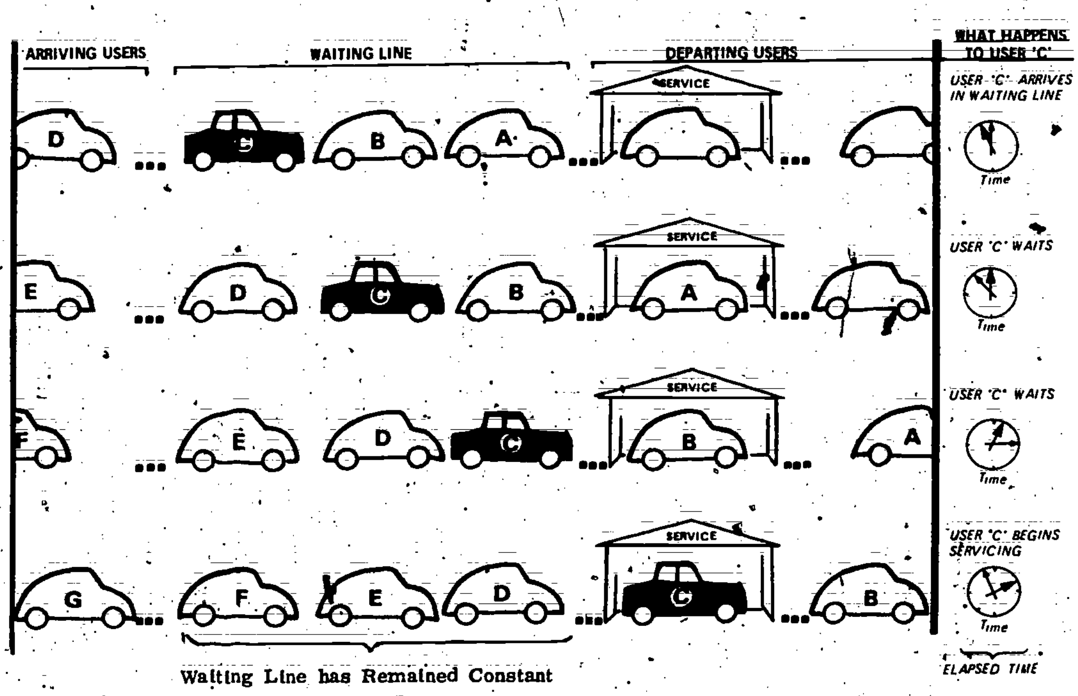


Figure 4-1.
Illustration of "Well-Behaved" Queueing System.

not leave it until he has been serviced. A customer is assumed not to get tired of waiting and leave before he is serviced.

6. *The arrival rate of customers must be less than the number of customers who can be serviced during the same unit of time.* This last condition is particularly important. If customers always arrive at a service facility at a rate *greater than* the service facility can handle them, the waiting line will become longer and longer; and there will be no assurance that every customer will eventually be serviced. Specifying that the arrival rate is less than the service rate assures us that all customers will eventually be serviced.

For a queueing system with just one service facility, this condition can be expressed simply as "the arrival rate of the customers must be less than the service rate"—or, symbolically, $A < S$.[1]

For queueing systems with multiple service facilities, if F is the number of facilities and S is the service rate for any one facility (assuming all have the same service rate), this condition can be expressed $A < F \cdot S$: that is, the arrival rate of customers is less than the total number of customers that can be handled by all service facilities. For example, if one purchasing clerk can process an average of 12 orders a day (S = 12 orders/day), two clerks (F = 2) could process $F \cdot S = 2 \cdot 12 = 24$ orders per day. The arrival rate of purchase orders, A, would then have to be less than 24 orders per day for this condition to be satisfied.

At this point, it may have occurred to you that we seem to have a paradox concerning arrival and service rates: if the number of arrivals at a service facility is always *less* than the number who can be serviced, how does a waiting line ever form? If we remember that the arrival rate and the service rate are *averages*, the paradox is easily resolved. Sometimes the number of arrivals may be higher than the average figure, and sometimes it may be lower. It is generally during the times when arrivals are greater than average that the waiting line is formed. Once a waiting line forms, the service facility falls behind in its work and may take some time to catch up with the arrivals; hence we will have a waiting line of customers which is assumed to stay at a fairly constant length.

[1] Since the sign <, means "less than," the possibility that $A = S$ is also excluded from any queueing problem.

SUMMARY OF CONDITIONS FOR QUEUEING THEORY PROBLEMS

The conditions a system must meet in order for queueing theory to be applicable are summarized below.

1. Arrivals of customers are random and independent.
2. Service times are independent.
3. Equilibrium system.
4. First come, first served.
5. Customers entering the queueing system do not leave until they are serviced.
6. Rate of arrivals must be less than the rate of service ($A < S$ for single facility systems; $A < F \cdot S$ for multiple facility systems).

These conditions have not been chosen arbitrarily; they are required by the mathematics used in queueing theory. The numerical results from queueing theory equations reflect the assumption that all these conditions are present in the system. Practical problems often violate one or more to a certain extent, as the realities of the everyday world do not often coincide with the strict conditions of mathematical theory. Small deviations are acceptable, but it is important to remember that the measures thus derived from queueing theory should be regarded as approximations to, rather than exact representations of, the real problem situation.

In the next section of this chapter, we apply the above terms and conditions to three general queueing problem situations and analyze each using queueing theory. The situations to be examined are:

1. *Infinite Population (30 or more), Single Service Facility.* An example of this system is a high school with one counselor to service all the students. We will examine situations in which one purchasing

clerk processes all purchase orders for a school, one switchboard operator handles all calls of a district office, and one repair facility is used to service a large population of school typewriters.

2. *Infinite Population, Multiple Service Facilities.* An example: five nurses (service facilities) serving all the students in a school represents an infinite-population multiple-service-facilities situation. The examples we will study are: an administrator considering having two operators handle all calls on the district's switchboard system; a school wanting to determine whether to purchase ten or fewer audio-tutorial listening stations for student use.

3. *Finite Population (less than 30) Single-Service Facility.* An example would be a situation in which each of 15 teachers is to review his annual evaluation with the principal. The situation we will examine involves a school wishing to determine if one science center is sufficient to serve eight teachers.

REVIEW

Exercises

1. Outline the conditions a queueing problem must satisfy before queueing theory can be applied. (Review text pages 94–96.)
2. What is the significance of these expressions?

   $A < S$
   $A < F \cdot S$ (Review text pages 96–97.)

3. A school district switchboard with one operator handles about 300 calls in one six-hour school day. The operator estimates that it takes about one minute to handle a typical call to greet the caller, determine whom the caller wishes to speak to, and transfer the call). Calls that come in when the operator is speaking to someone else are placed on hold until the operator is free.
   a. What is the service facility in this problem?
   b. Who are the customers?
   c. Is the source field infinite or finite?
   d. What is the arrival rate of customers? (Express in time units of minutes.)
   e. What is the service rate? (Also express in minutes.)
   f. If you were asked to determine if one operator was enough to run the switchboard efficiently, how would you answer? What criteria would you use in making your decision? (You are to use only the above data and your intuition at this point.)
4. A large city school district has 2 full-time psychometrists. Approximately 235 children must be given diagnostic tests each month. We will consider a month to consist of 20 school days, 6 hours a day. Tests are given on a one-to-one basis: that is, each psychometrist tests one child at a time. Records indicate that each child needs about one hour for testing.
   a. What are the service facilities in this problem?
   b. Who are the customers?
   c. Is the source field infinite or finite?
   d. What is the arrival rate of customers? (Express in time units of hours.)
   f. Does the district need to employ another psychometrist? What

criteria could you use to make your decision? (Again, use only the data given in the problem and your own judgment.)

APPLICATIONS OF QUEUEING THEORY

INFINITE POPULATION, SINGLE SERVICE FACILITY PROBLEMS

It is easy to visualize examples of a single service facility whose users are drawn from an infinite or essentially infinite population. The purchase order situation used earlier in this chapter is an example of a queueing situation with an infinite population of customers (orders) and a single service facility (the one purchasing clerk). The single clerk was in charge of processing all the purchase orders; the duties included making sure the purchase forms were properly completed and authorized, seeing which suppliers stocked the desired item, deciding which supplier received the order (according to price and speed of delivery), and finally placing the order by phone or mail. We will continue to use this example as we study the application of queueing theory to infinite population, single service facility problems.

Let's assume that recently your school administration has received complaints from teachers that they are having to do without certain materials. Some teachers have implied that there are too many purchase orders for one clerk to handle and it has been suggested that another purchasing clerk be hired. Before making a decision and presenting the case to the school board, assume that you, a school administrator, want to ascertain the extent of the problem. You would want to know such things as the average time an order waits to be processed after reaching the clerk's desk, the average number of orders waiting on the clerk's desk at any given time, and the relative amount of time the present clerk actually spends processing orders. With figures like these in hand, you could present an effective case before the board for hiring another clerk, if that were your final decision.

What you would need to do, then, is to calculate the quantities listed above. The formulas are not difficult, and we will go through each of them step by step. It is *not* necessary to memorize them. They have been included here to illustrate how queueing statistics are calculated and what information is necessary for these calculations. As an administrator you will in many cases be using computer programs to carry out queueing analyses. In the next section we introduce such a program and discuss its use in solving the purchase order problem, as an example. At this point, however, let's proceed with the hand calculations so that you can clearly grasp each step in the solution of this kind of problem.

At the outset, you can get a good idea of the *arrival rate* of purchase orders by examining the clerk's records for a substantial period of time. Let us use the data from pages 93–94 of this chapter, showing that purchase orders arrive at the rate of ten a day:

$$A = 10 \text{ orders/day}$$

and the clerk can usually get them processed at a *service rate* of 12 a day:

$$S = 12 \text{ orders/day.}$$

We now have all the information we need to analyze the problem using queueing theory:

Source field:	infinite (more than 30 orders)
Customers:	purchase orders
Service facilities:	one clerk (F = 1)
Arrival rate:	A = 10 orders/day
Service rate:	S = 12 orders/day

We will assume that this problem situation satisfies the conditions required for queueing theory: that the purchase orders arrive at random time intervals and independently of one another; the times necessary to process the orders are independent; the system is in equilibrium; the purchase orders are processed in the order in which they arrive; the orders do not leave the clerk's desk until they are processed; and the arrival rate is less than the service rate (A = 10 orders/day is less than S = 12 orders/day).

Let's start our analysis by determining how much of the clerk's time is occupied with processing orders and how much of his time is wasted. If the clerk is capable of processing 12 orders a day (service rate, S) but only receives 10 orders a day (arrival rate, A), then he is busy for only part of the day. The part of the day that he is busy is the fraction 10/12; the remainder must be free time for the clerk. This time, when the clerk is not processing orders, is called *idle time*. If the clerk is busy for 10/12 of a day, then he must be idle for 2/12 of the day, since

one full day − the part of the day the clerk is busy $= 1 - \frac{10}{12}$

$= \frac{12}{12} - \frac{10}{12}$

$= \frac{2}{12}$ part of the day the clerk is idle

The fractions expressing the amount of time the clerk is busy and idle can also be changed to percentages. The fraction 10/12 is the same as the decimal .83, so the clerk is busy for 83 percent of the day; 2/12 is decimal .17, so he is idle for 17 percent of the day.

Percentages can also be thought of as probabilities. To say the clerk is idle for 17 percent of the time is equivalent to saying that the probability that he is idle is .17. That is, if we look in on the clerk 100 times during the day, 17 of those times he will be idle. Our chances of seeing him idle are 17 out of 100, or .17.[2]

The above discussion shows the intuitive calculation for the probability that the clerk is idle. Now, let's state the general formula for this quantity. Remember that we are at the moment only working with an infinite population, single service facility problem. We will use the symbols P(0) (read P zero) to indicate the probability that the service facility is idle, i.e., the probability that zero customers are being serviced and zero customers are waiting for service. In other words, P(0) denotes the part of the time that the service facility is not servicing customers. The probability that the service facility is idle is calculated by this formula:

$$P(0) = 1 - \frac{A}{S}$$

where A is the average arrival rate and S is the average service rate.

For our purchase order problem, since A = 10 orders/day and S = 12 orders/day,

$$P(0) = 1 - \frac{A}{S} = 1 - \frac{10}{12} = \frac{2}{12} = .17$$

The next quantity we wish to find is how many orders, on the average, are waiting on the clerk's desk to be processed. We call this quantity the *expected number of customers waiting for service*, abbreviated in symbols E(W) (read E W). In queueing theory, "expected number" means the statistically estimated average number. The formula for E(W) is easy to use, but it does not have the simple, intuitive explanation that the formula for P(0) has. Consequently, you need only to understand the elements in the formula and to accept the equation itself on faith.

If A and S are respectively the average arrival and service rates, then E(W), *the expected number of customers waiting*, is:

$$E(W) = \frac{A^2}{S(S - A)}$$

For the purchase order problem, since A = 10 orders/day and S = 12 orders/day, the calculation of E(W) is:

$$E(W) = \frac{A^2}{S(S - A)} = \frac{(10)^2}{12(12 - 10)}$$

$$= \frac{100}{24} = 4.2 \text{ orders}$$

Note that this quantity *does* have units: E(W) = 4.2 *customers* or *orders*. That is, an average of 4.2 orders will be waiting on the clerk's desk at any given time; or, the average length of the waiting line of purchase orders is 4.2 orders. (This number does not include any order which the clerk is in the midst of processing.)

At this point, then, we know two important things about the system. First, there is a .17

[2] Probabilities can take values from 0 to 1, where a probability of 0 denotes an impossible event and a probability of 1 indicates a sure event.

probability that the clerk is idle; second, that the average length of the waiting line is 4.2 orders.

The final statistic of interest in this problem is *the average time an order has to wait in the line before the clerk begins to process it.* This quantity is abbreviated E(T) (read E T). The formula for E(T), the average waiting time of an arrival, is

$$E(T) = \frac{E(W)}{A}$$

where A is the arrival rate and E(W) is the average length of the waiting line.

In the purchase order problem, we know (E(W) = 4.2 orders and A = 10 orders/day; therefore,

$$E(T) = \frac{E(W)}{A} = \frac{4.2 \text{ orders}}{10 \text{ orders/day}} = 0.42 \text{ days}$$

The intuitive reasoning behind the derivation of this formula is as follows: Over long periods of time, we can speak of average arrival and service rates for a queueing system, and of an average waiting line length, which we have called E(W). Because of the steady arrival and service rates characteristic of a system in equilibrium, the waiting line length, E(W), always remains fairly constant; this means that orders must be leaving the clerk's desk at approximately the same rate as they are arriving. That is, the departure of the orders from the waiting line is at the same rate as their arrival. If the arrival rate is A, the departure rate is also A. When an order joins the waiting line on the clerk's desk, then, there are E(W) orders ahead of it, which are departing the waiting line at rate A. The last order to join the line will be processed after all orders in front of it have left the line. If 4.2 orders (E(W)) are on line and are departing at a rate of 10 orders/day (A), then an order which joins the line must wait .42 days = 4.2 orders divided by 10 orders/day = E(W)/A. (Note that the departure rate is *not* the same thing as the service rate, S. The simple illustration of Car C joining a waiting line and going through the service facility, as shown in Figure 4-1, again may help you visualize the elements of a queueing situation.)

We can now summarize all the information we have gathered about the purchase order situation:

Source field:	infinite
Customers:	purchase orders
Service facilities:	one clerk (F = 1)
Arrival rate:	A = 10 orders/day
Service rate:	S = 12 orders/day

Probability that service facility is idle = .17

$$P(0) = 1 - \frac{A}{S} = 1 - \frac{10}{12} = \frac{2}{12} = .17$$

Average length of waiting line = 4.2 orders

$$E(W) = \frac{A^2}{S(S-A)} = \frac{10^2}{12(12-10)} = \frac{100}{24} = 4.2 \text{ orders}$$

Average waiting time = .42 days

$$E(T) = \frac{E(W)}{A} = \frac{4.2 \text{ orders}}{10 \text{ orders/day}} = .42 \text{ days}$$

At this point, the administrator is in a position to examine the gathered data and make a decision. From the facts, he or she may conclude that any large stacking up of orders on the clerk's desk is fairly infrequent, since on the average only 4.2 orders (E(W)) are waiting on the clerk's desk at any given time. The average time an order waits to be processed is really quite small—less than half a day (E(T) = .42 days). Hence, the administrator may feel at this point that hiring another clerk is not justified; the present clerk seems to be handling orders fairly efficiently. The delay in getting materials is probably due more to inefficiency on the supplier's part than to the school district. Based on the observations of the current situation, the administrator now has several options open to him. He or she may instruct the clerk to work overtime when orders are heavy, or to order from fast delivery suppliers regardless of extra cost; or he may institute a system for teachers to order materials at the beginning of the school year. These suggestions are by no means the only possible ones. The administrator's ultimate decision will be based on which strategy he or she feels will best suit the circumstances and what specific constraints limit the decision. The school board, for example, may or may not be amenable to approving overtime pay for the purchasing clerk. Queueing theory has helped the administrator in this example to organize his available data and come to a quick appraisal of the initial suggestion of hiring another purchasing clerk.

Let's now look at the problem of the one switchboard operator for the school district, used in Exercise 3 on page 97. This is another example of an infinite population, single service facility queueing theory problem. The arrival rate of incoming calls to the district switchboard is 300 per day, and the switchboard operator works a six-hour day. It is estimated that the operator could handle about one call a minute. After expressing A = 300 calls per day as the equivalent figure of A = .83 calls per minute,[3] we have the following information:

Source field:	infinite
Customers:	incoming calls
Service facilities:	one switchboard operator (F = 1)
Arrival rate:	A = .83 calls/minute
Service rate:	S = 1 call/minute

(Again, note A < S.)

Then, the probability that the operator is idle, P(0), is:

$$P(0) = 1 - \frac{A}{S} = 1 - \frac{.83}{1}$$

$= .17$, or 17 percent of the time

The average number of incoming calls waiting to be answered (presumably on hold) is:

$$E(W) = \frac{A^2}{S(S-A)} = \frac{(.83)^2}{1(1-.83)} = \frac{.69}{.17} = 4.1 \text{ calls}$$

The average waiting time for a call is then

$$E(T) = \frac{E(W)}{A} = \frac{4.1 \text{ calls}}{.83 \text{ calls/minute}} = 4.9 \text{ minutes}$$

The statistic the administrator would probably feel is the most important is the waiting time of an incoming call, E(T). He or she may feel that almost five minutes (E(T) = 4.9 minutes) is actually quite a long time for a call to be held, especially when there may be an urgent message for one of the students in the district. In addition, the data used in this problem has not taken into account the outgoing long distance calls which the switchboard must also handle. Therefore, the delays before incoming calls are taken care of are probably greater than is indicated by the figure E(T).

At this point, the administrator probably would begin to seriously investigate the cost of adding another switchboard operator and would want to obtain estimates of what effect a second operator would have on waiting times for calls. The effect of adding another operator will be examined in the section in this chapter dealing with infinite populations and *multiple* service facilities, pages 108–113.

REVIEW

Terminology

1. IDLE TIME: time spent by facility not servicing customers.
2. P(O): probability that service facility is idle.
3. E(W): expected number of customers waiting for service.
4. E(T): average waiting time for a customer.

Exercises

1. In Exercise 3 on page 97, you were asked to analyze the district switchboard problem intuitively. Now that we have calculated P(O), E(W), and E(T), how is your first analysis affected? Can you use these data to improve your criteria for decision making in this problem?
2. Your public relations officer has made a request to you for an assistant. He reports he has had 210 telephone inquiries in the last month and averages about that many each month. (A school month will be assumed

[3] One day = 6 hours per day times 60 minutes per hour = 360 minutes. 300 calls/360 minutes = .83 calls/minute.

to be 20 days and 6 hours a day.) He also claims that it takes him half an hour to process each inquiry.

a. What is the service facility?
b. Who are the customers?
c. Is the source field finite or infinite?
d. What is the arrival rate (in terms of hours)?
e. What is the service rate (also in terms of hours)?
f. What is the probability that the service facility is idle?
g. What is the average length of the waiting line?
h. What is the average waiting time for customers?
i. If your school board has set down a policy that all public inquiries should be answered "promptly," what would your decision be regarding the new assistant for your public information officer? What criteria could you use?

SOLVING INFINITE POPULATION, SINGLE SERVICE FACILITY QUEUEING PROBLEMS WITH A COMPUTER

As can be seen from the examples given thus far, the application of the queueing theory formulas, even in relatively simple problems, can involve some tedious arithmetic computation. Use of a computer can greatly facilitate an administrator's ability to gather statistics on queueing problems and to simulate potential changes in the system. Hence, at this point, we will move from manual to computer-based calculations of queueing statistics.

A program to solve the queueing theory equations has been written in BASIC (Beginner's All-Purpose Symbolic Instruction Code). The program, named QUEUE, requires a minimal amount of input. To solve a problem with it, you need to know only the basic information about the problem that you have already learned to identify:

1. The size of the source field (finite or infinite);
2. The number of service facilities, F;
3. The average arrival rate of customers, A;
4. The average service rate, S.

QUEUE will calculate and print out the following quantities on the basis of the data you enter:

P(0)	probability that the facility is idle
P(N)	probability that N users are being serviced or waiting
P(M)	probability that all persons in source field are being serviced or waiting (P(M) = 0 for an infinite source population)
E(N)	expected number being serviced or waiting
E(W)	expected number waiting
E(T)	expected waiting time
P(N > L)	probability that the number of customers being serviced or waiting exceeds L

The quantities P(0), E(W), and E(T) have already been explained in this chapter. P(N) and P(N > L), which will be illustrated in subsequent problems, are useful when the length of the waiting line is important to a particular problem. E(N), the expected number being serviced or waiting, differs from E(W), the expected number waiting, in that E(N) includes customers in the process of being serviced as well as customers in the waiting line; E(W) includes only the average number of customers in the waiting line. The quantity P(M) applies to problems dealing with finite populations and will be discussed when we study the problems on pages 115–117.

To see how QUEUE works, let's take the purchase order problem which was analyzed on pages 93–94 and solve it using the QUEUE program. Recall that we had the following information for this problem:

```
GET-QUEUE
RUN
QUEUE

ENTER SOURCE FIELD
      O=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?1
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?10
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?12

PROBABILITY THAT THE FACILITY IS IDLE: P(O) =  .16667
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 5
EXPECTED NUMBER WAITING: E(W) =  4.16667
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =  .41667

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
      SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
      ENTER -1
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
      USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
      ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?-1

ENTER SOURCE FIELD
      O=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?-1

DONE
```

Figure 4-2.
Sample Run of QUEUE, Purchase Order Problem.

Source field:	infinite
Customers:	purchase orders
Service facilities:	one clerk (F = 1)
Arrival rate:	A = 10 orders/day
Service rate:	S = 12 orders/day

This basic information is all we need. A complete listing of the computer run of QUEUE for this problem is shown in Figure 4-2, but let's start here by looking at the program run one step at a time.

Since the first thing to be done is to tell the computer that you want to run the program called QUEUE, you would type the appropriate statements on the terminal.[4] For example:

```
GET-QUEUE
RUN
```

These statements have been underlined to indicate that they are typed by the user (you), not the computer. The underlining will, of course, not appear when you actually type entries on the teletypewriter.

Next, the computer will ask for the type of population by printing the following:

```
ENTER SOURCE FIELD
      O=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?0
```

(We have underlined the zero to identify it as an item the user types.)

If we had been dealing with a finite population, we would have typed in the actual size of the population (denoted by M) instead of 0. If we did not wish to run the program any further, we would enter -1 and the program would stop. The program will next ask for the number of service facilities. Since this problem has only one (a purchasing clerk), a 1 would be entered:

```
F - NUMBER OF SERVICE FACILITIES
?1
```

The program will next ask for the average number of arrivals per unit time. In our prob-

[4] If you have never used a terminal to run a computer program, consult your instructor. It should also be noted that statements required by your computer to access the program QUEUE and begin its execution may be different from those used in this text (GET-QUEUE and RUN); be sure you know what statements are appropriate for your computer.

lem, ten orders per day arrived, so a 10 should be entered as illustrated:

```
A = AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?10
```

The last figure the program will ask for is the average number of customers served by the facility per unit of time. In our example, the clerk can process an average of 12 orders per day, so a 12 should be entered after the question mark. Remember that the units of time for arrival rate A and service rate S must be the same (here, both are one day).

```
S = AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?12
```

We have now input all the information the program needs to calculate the queueing statistics. After the calculations are made, the program will print out the following statistics:

```
PROBABILITY THAT THE FACILITY IS IDLE: P(0) = .16667
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 5
EXPECTED NUMBER WAITING: E(W) = 4.16667
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) = .41667
```

These figures, with the addition of E(N), are the same ones obtained on page 101 using manual calculation: the purchasing clerk is idle about 17 percent of the time (P(0) = .17, (which is .16667 rounded to two places); an average of 4.2 orders will be waiting to be processed at any given time (E(W) = 4.2); and the expected waiting time of a customer, or the average time from when an order reaches the clerk's desk to when the clerk begins to process it, is .42 days, or less than half a day (E(T) = .42).

At this point, the program will ask you if you wish to find P(N) for some N, that is, the probability that exactly N customers are in the waiting line or being serviced. For the time being, we will reject this option; therefore, a -1 would be entered in response.

```
IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
    USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
    ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?-1
```

The final option you will be given is to enter some number, L, to find P(N > L), the probability that *more than* L customers are in the waiting line. Again, let's reject that option for now; it will be illustrated in a la[illegible] example.

```
IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
    SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
    ENTER -1
?-1
```

At this point, the program is finished with the data you originally gave it; it is now ready for new data, and will give you the following instructions, which appeared before, in the beginning of the program:

```
ENTER SOURCE FIELD
    0=INFINITE POPULATION
    N = FINITE POPULATION OF SIZE N
    -1 TO QUIT
?
```

If you have another problem to solve, merely proceed as outlined above, first specifying your source field and so forth. If you have no more data, enter -1 in response to the question mark and the computer will type:

```
DONE
```

To see the entire run we just made, read through the printout illustrated in Figure 4-2.

It is important to remember that the computer program for queueing theory will give you data, not make your decision. Based on the size of the source field, the number of service facilities, the arrival rate, and the service rate, the program will tell you the probability of the facility being idle, the expected waiting line length, the expected waiting time, and the probabilities of certain numbers of customers waiting in line. It will *not* tell you whether the expected waiting line is too long; whether service facility is idle too much of the time; whether you can afford another service facility; or whether the waiting time for a customer is too long. These are decisions you must ultimately make. The data from queueing theory will often aid you greatly in making your decision; but queueing theory alone will not make the decisions.

REVIEW

Exercises

1. What basic information do we need to enter into the computer program QUEUE? (Review text pages 102-103.)

2. What statistics does QUEUE calculate? (Review text page 102.)
3. How does the statistic E(N) differ from E(W)? (Review text page 102.)
4. What is the number that must be entered to reject the options of finding P(N) and P(N > L)? (Review text page 104.)

Let us now consider a queueing problem, of a somewhat different type but again one involving a single service facility and an infinite population. Suppose that a school has a large number of typewriters (greater than 30) used for instructional purposes. They also have eight spare typewriters which are brought into use when any of the regular typewriters break down and have to be sent to a local repair shop. Records indicate that typewriters break down on the average of one every ten days and that it takes eight days for the repair shop to service a typewriter and return it to the school. The crucial concern here is the possible situation of more than eight typewriters being out for repair: the class would be short one or more typewriters, so at least one student would be without a typewriter. Thus, it is important to know how often this situation might occur. The quantity which will be of special interest here is P(N > 8), the probability that the number of typewriters waiting to be serviced or in the process of being serviced (N) is greater than the number of spare typewriters available (8).[5]

To analyze this problem using QUEUE, we first need to calculate arrival and service rates. Since one typewriter is sent for repair on the average of every ten days, the number of typewriters sent per day is 1/10; therefore, A = 0.1 typewriters/day. The service period is eight days, so the average service rate is 1/8 typewriters per day; S = .125 typewriters per day. In summary, then, we have:

Source field: infinite
Customers: typewriters needing repair
Service facilities: one repair shop (F = 1)
Arrival rate: A = .1 typewriters/day
Service rate: S = .125 typewriters/day

The complete computer listing for this problem is Figure 4-3, page 106. The appropriate responses by the user would be 0 to indicate an infinite source field, 1 to indicate one service facility, .1 to denote the arrival rate of .1 typewriters/day, and .125 for a service rate of .125 typewriters a day:

```
ENTER SOURCE FIELD
    0=INFINITE POPULATION
    N = FINITE POPULATION OF SIZE N
    -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?1
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.10
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?.125
```

The statistics which QUEUE then calculates will be:

```
PROBABILITY THAT THE FACILITY IS IDLE - P(O) =  .2
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING - E(N) =  4
EXPECTED NUMBER WAITING - E(W) =  3.2
EXPECTED WAITING TIME OF AN ARRIVAL - E(T) =  32.
```

That is, the repair shop is idle 20 percent of the time (P(O) = .20); an average of 4 typewriters are being serviced or waiting to be serviced (E(W) = 3.2); and the average waiting time for a typewriter until repairs are begun is 32 days (E(T) = 32). Since our primary interest in this problem is finding P(N > 8)—the probability that more than 8 typewriters will be tied up in the repair shop—we will first reject the option of finding P(N) for some N by entering -1 to this option:

```
IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
    SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
    ENTER -1
?-1
```

Then, when the computer presents the next option, we will enter 8 to specify L:

```
IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
    USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
    ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?8
```

The computer will respond immediately with,

```
P(N > 8     )   .13422
ENTER ANOTHER L OR -1 TO QUIT
?
```

Let's hold off on our response to ? for the time being. Now we know that 13 percent of the time (P(N) > 8) = .13), more than 8 typewriters will be tied up in the repair shop; that is, 13 percent of the time, at least one student will be without a typewriter.

As an administrator reviewing the above data,

[5] Remember, the sign > means "greater than."

```
GET-QUEUE
RUN
QUEUE


ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?1
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.1
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?.125

PROBABILITY THAT THE FACILITY IS IDLE: P(0) = .2
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 4
EXPECTED NUMBER WAITING: E(W) = 3.2
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   32.

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?8
P(N > 8     ) =  .13422
ENTER ANOTHER L OR -1 TO QUIT
?9
P(N > 9     ) =  .10737
ENTER ANOTHER L OR -1 TO QUIT
?10
P(N > 10    ) =  .0859
ENTER ANOTHER L OR -1 TO QUIT
?11
P(N > 11    ) =  .06872
ENTER ANOTHER L OR -1 TO QUIT
?12
P(N > 12    ) =  .05498
ENTER ANOTHER L OR -1 TO QUIT
?13
P(N > 13    ) =  .04398
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?-1

DONE
```

Figure 4-3.
Sample Run of QUEUE Typewriter Problem.

you promptly determine that students are without typewriters too much of the time. You would be likely to ask: How many more typewriters are needed to reduce the probability of one or more students being without a typewriter to an acceptable level? The answer to this question depends on how much of the time you are willing to have students without typewriters. For example, assume that you want students to have typewriters available 95 percent of the time. That is equivalent to stating that you want enough typewriters so that the probability of one or more students being without a typewriter is less than .05. Suppose you decide to see whether purchasing one more spare typewriter will reduce the typewriterless situation to .05 probability. Then, you will want to know what the probability is that more than nine typewriters are tied up in the repair shop. At this point in the computer run, you could have entered 9 in response to the above option to "ENTER ANOTHER L OR -1 to QUIT," and obtained:

```
?9
P(N > 9     ) = .10737
ENTER ANOTHER L OR -1 TO QUIT
?
```

That is, more than nine typewriters will be tied up in the repair shop about 11 percent of the time. The probability has decreased from $P(N > 8) = 13$ percent but it is still not 5 percent, which you decided was an acceptable level. So, you must try $P(N > 10)$, $P(N > 11)$, and so on, until the probability is 5 percent or less. The complete sequence of requests and responses for $P(N > L)$ would be as follows:

```
IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
    USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L.
    ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?8
P(N> 8    ) =    .13422
ENTER ANOTHER L OR -1 TO QUIT
?9
P(N> 9    ) =    .10737
ENTER ANOTHER L OR -1 TO QUIT
?10
P(N> 10   ) =    .0859
ENTER ANOTHER L OR -1 TO QUIT
?11
P(N> 11   ) =    .06872
ENTER ANOTHER L OR -1 TO QUIT
?12
P(N> 12   ) =    .05498
ENTER ANOTHER L OR -1 TO QUIT
?13
P(N> 13   ) =    .04398
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
    0=INFINITE POPULATION
    M = FINITE POPULATION OF SIZE M
    -1 TO QUIT
?-1

DONE
```

The probability we are interested in, .05, falls between .05498 and .04398; that is, the value of L which makes $P(N > L) = .05$ is somewhere between $L = 12$ and $L = 13$. Since we cannot buy fractions of typewriters, we will choose $L = 13$. That is, having 13 spare typewriters will assure us that students are without typewriters less than 5 percent of the time.

Thus, in order to ensure theoretically that all students will have typewriters 95 percent of the time, an additional five machines should be purchased, since the school originally had 8 spare typewriters and we want to have a total of 13.

Of course, buying additional spare typewriters was not the only possible solution to the typewriter problem. We could have investigated ways of decreasing the service time S (for example, by using more than one repair shop), or we could have considered buying typewriters of better quality so the arrival rate A was smaller (i.e. rate at which typewriters broke down and arrived at the service facility for repairs). The final decision would then be based on which modification of the queueing system was best in terms of cost, time needed to make the changes, and resulting probabilities of students being without typewriters.

These examples should give you some idea of how queueing theory can be applied to situations where one has a waiting line formed by arrivals from an infinite population of customers and only one facility to serve them. You need only to determine two quantities, the average number of arrivals per unit time A and the average number serviced per unit time S in order to calculate several useful characteristics of the queueing system you are interested in.

REVIEW

Exercises

1. (a) Verify our manual calculations for the switchboard problem, page 101, using QUEUE.
   (b) Suppose six outside lines are available on the switchboard. When more than six callers are trying to reach the school district, they will get a busy signal. Find $P(N > 6)$ for this problem to see how often callers to the school district get a busy signal. Would you be satisfied with this answer? If not, what actions would you take?
2. Your school district has one man in charge of refinishing and repairing students' desks. He claims he is overloaded with work, and also needs a lot more storage space for desks waiting for repair. His records show he repairs 135 desks a month, and the average desk needs about 1 hour's work. Assume the repairman works 20 days a month, 8 hours a day.

Choosing the time units you feel are appropriate to the problem, find, using QUEUE:

(a) the probability that the repairman is idle;
(b) the average number of desks waiting for repair;
(c) the average time a desk waits for repair;
(d) the probability that more than 20 desks are waiting for repairs or are being repaired.

What criteria could you use in making your final decisions?

INFINITE POPULATION, MULTIPLE SERVICE FACILITIES PROBLEMS AND THEIR COMPUTER SOLUTIONS

On page 101, we examined the situation in which one switchboard operator serviced the needs of an entire school district. If as a result of this analysis the administrator decided that it was desirable to add another operator to the staff, he might wish to examine the effects this additional operator would have on telephone call waiting time, operator idle time, and the average number of incoming calls waiting to be handled. The problem now involves an infinite population and *multiple service facilities*. Two conditions (beyond those required for the single service facility case) are required for this type of problem:

1. Each facility must have the same service rate.
2. Only a single waiting line may be formed.

To elaborate briefly on the first condition, recall from page 93 that service rate was defined as the average number of customers which can be served by *one* facility per unit of time. With multiple facilities, we still consider service rate to be the rate of *one* of the facilities. We therefore must assume that all service facilities work at the same rate. The second condition specifies that customers stand in one waiting line for all of the facilities. As a facility becomes available, the next person in line goes to that facility for servicing. In a situation with four service facilities, for example, we still only have one waiting line, not four.

We will deal with queueing problems involving multiple facilities only on the computer, since the formulas become very complex and have no obvious intuitive explanations.[6] Multiple-facility problems are quite easy to handle using QUEUE, however, since the user needs only to enter the appropriate number of facilities for F along with the other routine data.

Suppose an administrator wanted to see what specific effects the addition of another operator would have on the district switchboard system. Using the data from the original switchboard problem, but changing the number of facilities to two, we have:

Source field: infinite
Customers: incoming calls
Service facilities: two operators (F = 2)
Arrival rate: A = .83 calls/minute
Service rate: S = 1 call/minute

The appropriate user responses to the data requests from QUEUE for this multiple-facility situation would be:

```
ENTER SOURCE FIELD
   0=INFINITE POPULATION
   M = FINITE POPULATION OF SIZE M
   -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?2
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.83
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?1.0
```

Notice that A and S remain the same as in the problem involving only one operator; service facilities are assumed to operate at the same rates and calls to come into the switchboard with the same frequency.

The computer program will respond with values for P(O), E(N), E(W), and E(T), as shown in the computer listing in Figure 4-4, which combines the results for the one-operator and

[6] Formulas for multiple-facility statistics are included at the end of this chapter for your reference.

```
GET-QUEUE
RUN
QUEUE


ENTER SOURCE FIELD
     O=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?1
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.83
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?1.0

PROBABILITY THAT THE FACILITY IS IDLE: P(0) =  .17
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 4.88235
EXPECTED NUMBER WAITING: E(W) =  4.05235
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   4.88235

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?-1

ENTER SOURCE FIELD
     O=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?2
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.83
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?1.0

PROBABILITY THAT THE FACILITY IS IDLE: P(0) =  .41343
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 1.00269
EXPECTED NUMBER WAITING: E(W) =  .17269
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   .20806

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?-1

ENTER SOURCE FIELD
     O=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?-1

DONE
```

Figure 4-4.
Run of QUEUE—Switchboard Operator Problems.

two-operator problems. Let us compare the statistics for the one-operator and two-operator problems on page 110. This table compares the pertinent statistics between the one- and two-operator systems. It shows the dramatic improvements in the average number of calls waiting and the average waiting time could be anticipated by adding the second operator. These improvements are offset, of course, by the added salary and overhead expense of the second operator and the increase in idle time, though "idle time" in this instance refers only to the amount of time not spent in answering and transferring incoming calls. Presumably

	One Operator	Two Operators	Change
Time all facilities are idle: P(0)	17%	41%	+24%
Expected number of calls waiting: E(W)	4.1	.2	-3.9 calls
Expected waiting time per call: E(T)	4.9	.2	-4.7 minutes

some of it could be utilized for other tasks, such as placing outgoing long distance calls or doing clerical tasks like folding memos or stuffing envelopes.

An administrator who approached the school board to request the hiring of another switchboard operator would have firm evidence of the probable effectiveness of such an addition if he presented the data in the above table. Of course, the final decision of whether to hire another operator will ultimately depend on the school budget, and there is not much the administrator can do if the funds for another operator are just not available.

You can probably appreciate the usefulness of queueing theory much more now that you have seen how it allows you to simulate changes in the queueing system and observe the resulting effects quickly and at little cost. Using this strategy, you can easily determine how many facilities are ideal for a particular situation, then make the decision which will result in the best possible operation of the system at the least possible cost.

The number of facilities, however, is not the only part of a queueing problem that could be varied. A school administrator could, for example, predict changes in the arrival rate for incoming calls, based on projected growth figures for his district, and then—using queueing theory—obtain estimates on the future waiting times for calls if he continued to use one switchboard operator. It might also be realistic to alter the service rate fo[illegible]articular problem and observe the results. I[illegible] words, the number of service facilities, [illegible]rrival rate, and the service rate may all [illegible]n as controllable variables in a queueing pro[illegible]

It is in these are[illegible]simulation and hypothesizing system chang[illegible] that queueing theory is most useful. In some of the problems we have analyzed thus far, it would have been possible to obtain such statistics as average waiting times for customers by directly observing the system in operation over a long period of time; but this direct observation would have been a far more time-consuming way of obtaining the data. In addition, it is impossible to "observe" the theoretical effect of adding another facility without a tool like queueing theory.

Still, it is important to remember that queueing theory necessarily gives only approximations to real-life situations. The further the real-life queueing system deviates from the basic conditions noted on pages 94-96 the more the resulting queueing statistics will vary from what will actually occur.

Let's consider another example of a multiple service facility problem with an infinite source population. Suppose that a school has an audio-tutorial laboratory, which students in several classes use. They come to the laboratory at random times throughout the day and listen to audiotapes as part of their classwork. There are ten listening stations, which have been leased on a trial basis. An average of twelve students an hour come to the laboratory, and an average of two students per hour can be served by each listening station. The school is now at a point where it wants to purchase the listening stations instead of leasing them. The question you have been given to answer is, Does the school really need to purchase all ten stations, or would a smaller number do just as well?

We have the following data for this problem:

Source field: infinite
Customers: students
Service facilities: ten listening stations (F = 10)
Arrival rate: A = 12 students/hour
Service rate: S = 2 students/hour

It is in approaching this type of problem that an administrator can really appreciate queueing theory; he can "play" with the data to simulate proposed changes in the system and get results quickly and at a very small cost.

After making an initial run of the program QUEUE with the above data, assume you obtain the following information. (Refer to the complete computer listing for this problem in Figure 4-5, pages 112-113.)

```
PROBABILITY THAT THE FACILITY IS IDLE: P(O) = .00243
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 6.15195
EXPECTED NUMBER WAITING: E(W) = .15195
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) = .01266
```

What do these data initially indicate to the administrator?

- *All* listening stations are almost never idle at the same time (P(O) = .00243).
- But the *average* number of listening stations in use at any one time is only about 6. (E(N) = 6.2; i.e. 6.2 students are waiting or being serviced at any one time. A student would not be waiting if a listening station were available. Therefore, E(N) = 6.2 indicates that, on the average, about 6 students are using the facilities at any given time.)
- Almost no students are ever waiting in line to use the audio-tutorial laboratory (E(W) = .15195).
- The waiting time to use a listening station is almost nil (E(T) = .01266).

At this point, you may have cause to suspect that having ten listening stations is a luxury. A good question you may ask here is, "What is the probability that *all ten* listening stations are busy? To get an estimate of the probability that all stations are busy, you will need to obtain two quantities: P(10) and P(N > 10). P(10) will give the probability that ten students are using the facilities and no students are waiting; P(N > 10) is the probability that ten students are using the facilities and there is someone waiting (i.e. more than ten customers are being serviced or are waiting). The sum of these two quantities, then, will give the probability that all ten listening stations are busy, with or without any students waiting.

From QUEUE, we find:

$$P(10) = .04$$
$$P(N > 10) = .05$$

Therefore:

$$P(10) + P(N > 10) = .04 + .05 = .09$$

That is, all the listening stations are in use only 9 percent of the time!

This result plus the previous data may encourage you to simulate the queueing system using fewer than ten listening facilities and to observe the effect of having a smaller number of facilities on facility idle time, student waiting time, and waiting line length. If you tried to simulate the system with fewer than seven facilities, you would receive the message CONGESTED SYSTEM and the computer program would stop. The reason for this message is that having fewer than F = 7 facilities violates condition 6 on page 00: for any multiple facility queueing system, it must be the case that A < F · S. Since A = 12 and S = 2, then F must be 7 or more.

After your simulations have been completed (see Figure 4-5, pages 112-113), you would summarize your data as follows:

	No. of Listening Stations, F			
	7	*8*	*9*	*10*
Probability that all facilities are idle P(0)	0	0	0	0
Average number of facilities in use at any one time F(N)	all	7.1	6.4	6.2
Expected students waiting E(W)	3.7	1.1	0.4	0.2
Expected waiting time (hours) E(T)	0.3	0.1	0.0	0.0
Probability that all stations in system are in use: P(F) + P(N > F)	.20	.20	.15	.09

Based on the above data, you could now justify the decision to purchase eight listening stations: you could then be fairly certain that the stations would receive maximum possible use. At the same time, the students would not be inconvenienced by having only seven listening stations. The expected waiting time for any student using eight facilities is only .1 hour, or 6 minutes,

```
GET-QUEUE
RUN
QUEUE

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F = NUMBER OF SERVICE FACILITIES
?10
A = AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?12
S = AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?2

PROBABILITY THAT THE FACILITY IS IDLE: P(O) =  .00243
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 6.15195
EXPECTED NUMBER WAITING: E(W) =  .15195
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   .01266

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?10
P( 10   )=   .04058
ENTER ANOTHER N OR -1 TO QUIT
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.   IF NOT, ENTER -1
?10
P(N> 10    ) =   .04662
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F = NUMBER OF SERVICE FACILITIES
?7
A = AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?12
S = AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?2

PROBABILITY THAT THE FACILITY IS IDLE: P(O) =  .00158
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 9.68298
EXPECTED NUMBER WAITING: E(W) =  3.68298
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   .30692

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?7
P( 7    )=    .08769
ENTER ANOTHER N OR -1 TO QUIT
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.   IF NOT, ENTER -1
?7
P(N> 7    ) =    .10523
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F = NUMBER OF SERVICE FACILITIES
?6
A = AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?12
S = AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?2
```

continued

Figure 4-6.
Audio-Tutorial Laboratory Problem.

```
PROBABILITY THAT THE FACILITY IS IDLE: P(0) = .00214
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 7.07094
EXPECTED NUMBER WAITING: E(W) = 1.07094
EXPECTED WAITING TIME OF AN ARRIVAL: E(T) =   .08985

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?8
P( 8    )=       .08985
ENTER ANOTHER N OR -1 TO QUIT
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?8
P(N> 8     ) =    .10709
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?0
F - NUMBER OF SERVICE FACILITIES
?9
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?12
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?8

PROBABILITY THAT THE FACILITY IS IDLE: P(0) =  .00235
EXPECTED NUMBER EITHER BEING SERVICED OR WAITING: E(N) = 6.39196
EXPECTED NUMBER WAITING: E(W) =  .39196
EXPECTED WAITINO TIME OF AN ARRIVAL: E(T) =   .03866

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?9
P( 9    )=       .06533
ENTER ANOTHER N OR -1 TO QUIT
?-1

IF YOU WISH P(N>L), THE PROBABILITY THAT THE NUMBER OF
     USERS BEING SERVICED OR WAITING EXCEEDS SOME NUMBER L,
     ENTER THE NUMBER FOR L.  IF NOT, ENTER -1
?9
P(N> 9     ) =    .07839
ENTER ANOTHER L OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
     0=INFINITE POPULATION
     M = FINITE POPULATION OF SIZE M
     -1 TO QUIT
?-1

DONE
```

Figure 4-5 Continued

and this time could be reduced by instituting some such procedure as a sign-up sheet, instead of having students come to the laboratory on a random basis.

Other considerations you, as the administrator, might want to weigh before making a final decision, however, are:

1. How frequently do stations need repairs? (Would it be wise to buy eight stations and have one in reserve?)
2. Is the arrival rate going to increase significantly in the next few years as a result of school population growth?
3. Will there be a large increase in the number of students taking courses which will require use of the listening stations?

It would be entirely possible to answer even these questions using queueing theory, if you could obtain the relevant data (rate of breakdowns, increased arrival rates, and so forth).

REVIEW

Exercises

1. What assumptions are necessary when working with an infinite population, multiple service facility queueing system, which are not required when working with infinite population, single facility systems? (Review text page 108.)
2. To obtain the probability that one facility is idle and the rest are in service in a problem involving five service facilities, which of the following quantities would you request?
   a. P(5) c. P(1)
   b. P(0) d. P(4)
3. Which of the following expressions would give you the probability that all five facilities in a queueing system (F = 5) are busy, with or without customers waiting?
   a. E(5) c. P(5) + P(N > 5)
   b. P(5) d. P(N > 5)
4. Based on the examples worked out in this chapter, as you increase the number of service facilities in an infinite population queueing theory problem while holding the arrival and service rates constant, what is the effect on:
   a. the expected length of the waiting line, E(W)?
   b. the expected waiting time, E(T)?
   That is, do E(W) and E(T) increase, decrease, or stay the same?
5. Analyze the school psychometrist exercise on pages 97–98 using the computer program QUEUE, and compare the new information you get from QUEUE with your initial intuitive analysis. What new information do you have? What criteria might you use in making a final decision?
6. Use QUEUE to answer the following question:
   A school district has 30 buildings requiring electrical repair work on a random basis. On the average, 1.5 buildings per week require such electrical work. How many electricians are needed if each electrician can handle 1.7 buildings per week?
7. Use QUEUE to analyze the following problem:
   Suppose you have been put in charge of the one-day kindergarten registration for your large school district. Based on last year's figures, an average of 94 mothers arrived per hour to register their children, and it took a worker about 12 minutes to complete registration for each child. That is, five registrations could be processed per hour by any one registration worker.
   You plan to recruit some teachers as registration workers, and you can pay them $20 each for their day's work. The question is, How many teachers should be hired to conduct the registration? There are several criteria you might consider in making your decision:
   (a) You don't want the mothers to have to wait in line too long;
   (b) You don't want to hire more teachers than are absolutely necessary (i.e. you want to keep costs down).
   (c) You want to be sure that enough teachers are used so all registrations can be completed.

FINITE POPULATION, SINGLE SERVICE FACILITY PROBLEMS AND THEIR COMPUTER SOLUTIONS

Up to this point we have considered only problems involving potential customers from an infinite or very large finite source field or population (30 or more). In actual queueing problems, we frequently have only a small number of potential customers. Examples of such populations (< 30) would be the fifth grade teachers in a single building; the buildings in a school district; the teacher aides in a single school. In addition, the queueing theory we have considered so far has assumed that the arrival of customers follows a completely random pattern—that they must arrive independently from an infinite source field. With finite populations, these assumptions are not satisfied.

Different equations must be used to take into account the peculiarities of a small source population. These are given for your reference in the "Summary of Equations" section on page 119.

To use QUEUE for such problems, the only change we need to make in our procedures is to enter the exact size of the population in response to the program's request for "source field." For example, if we were analyzing a queueing problem where the population was 8 customers, the correct response would be:

```
ENTER SOURCE FIELD
      0=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?8
```

In addition, there are two differences in the information QUEUE will give us:

1. Instead of the quantity E(N), we will be given P(M), the probability that *all* potential users in the source are being serviced or are waiting.
2. We will not have the option of obtaining P(N > L) for some L. But we *can* calculate probabilities of more than some number of customers (L) being serviced or waiting. For example, suppose we had a population of five customers and wanted to know the probability that more than three of them were being serviced or waiting. The probability that *more than three* customers are being serviced or waiting is the probability that four customers are being serviced or waiting plus the probability that five customers are being serviced or waiting, or P(4) + P(5).

Let's look at an example of a finite population single service facility queueing problem:

> Eight teachers use a portable science center. Each requires it an average of 18 times per school year of 180 days. Each time the science center is used, it is for an average of one day. Is one science center enough to serve the teachers satisfactorily?

To calculate the arrival rate for this problem, we know eight teachers each need the science center 18 times in a school year; therefore, the science center is used 8 × 18 = 144 times a year. The school year is 180 days, so:

A = 144 times/180 days
  = .8 times (or teachers)/day

Since the science center can service one teacher a day,

S = 1 time (or teacher)/day

The information we have, then, is

Source field: finite (M = 8)
Customers: eight teachers
Service facilities: one science center (F = 1)
Arrival rate: A = .8 teachers/day
Service rate: S = 1 teacher/day

The complete computer run for this problem is shown in Figure 4-6 on page 116. The most important quantity we would want to examine is probably E(T), the expected waiting time, which is about six days (E(T) = 5.8). A decision to add another science center would depend on the judgment of whether a six-day wait is too long.

```
GET-QUEUE
RUN
QUEUE


ENTER SOURCE FIELD
      0=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?8
F - NUMBER OF SERVICE FACILITIES
?1
A - AVERAGE NUMBER OF ARRIVALS PER UNIT TIME
?.8
S - AVERAGE NUMBER OF CUSTOMERS SERVED PER UNIT TIME
?1

PROBABILITY THAT THE FACILITY IS IDLE - P(0) =  .00004

PROBABILITY OF ALL CUSTOMERS WAITING OR BEING SERVICED
          P(M) = .28651
EXPECTED NUMBER WAITING - E(W) =  5.7501
EXPECTED WAITING TIME OF AN ARRIVAL - E(T) =  5.75034

IF YOU WISH P(N), THE PROBABILITY THAT N USERS ARE BEING
     SERVICED OR WAITING, ENTER THE NUMBER FOR N.  IF NOT,
     ENTER -1
?0
P( 0    ) =      .00004
ENTER ANOTHER N OR -1 TO QUIT
?1
P( 1    ) =      .00027
ENTER ANOTHER N OR -1 TO QUIT
?2
P( 2    ) =      .00158
ENTER ANOTHER N OR -1 TO QUIT
?3
P( 3    ) =      .00789
ENTER ANOTHER N OR -1 TO QUIT
?4
P( 4    ) =      .02914
ENTER ANOTHER N OR -1 TO QUIT
?5
P( 5    ) =      .09326
ENTER ANOTHER N OR -1 TO QUIT
?6
P( 6    ) =      .22383
ENTER ANOTHER N OR -1 TO QUIT
?7
P( 7    ) =      .35813
ENTER ANOTHER N OR -1 TO QUIT
?8
P( 8    ) =      .28651
ENTER ANOTHER N OR -1 TO QUIT
?-1

ENTER SOURCE FIELD
      0=INFINITE POPULATION
      M = FINITE POPULATION OF SIZE M
      -1 TO QUIT
?-1

DONE
```

Figure 4-6.
Science Center Problem.

Notice in the second half of the computer run in Figure 4-6 that the quantities P(0), P(1), P(2), P(3), P(4), P(5), P(6), P(7), and P(8) have all been requested. P(0) is, of course, the probability that no customers are being serviced or waiting or, equivalently, the probability that the facility is idle, which was already calculated in the first half of the run. P(1) is the probability that one customer is being serviced or waiting for service—i.e. the probability that one customer is being serviced and none are waiting. P(2) would then be the probability that one customer is being serviced and one is waiting; P(3), that one is being serviced and two are waiting; and so on. Notice that P(8) is the same as P(M). Both quantities are the probability that all potential customers in the source field are being serviced or waiting.

There is one major difficulty in dealing with queueing problems involving finite populations:

we cannot test alternative hypotheses about increasing the number of service facilities. This restriction results from the limits of the mathematical theory for queueing systems which has so far been developed. QUEUE will give an error message if we attempt to rerun the science center problem using two service facilities instead of one. So, we are limited to conjecturing about the exact impact of increasing the number of service facilities to some number greater than one. It is generally true that we can get only partial guidance from the solutions to finite population, single service facility problems. Thus these equations, though they are applicable in a number of situations, are not as useful as the equations for an infinite population with single or multiple service facilities.

REVIEW

Terminology

1. P(M): probability that all potential users in the source field are being serviced or are waiting.

Exercise

1. Use QUEUE to analyze the following problem:
   A school district has 25 school buses; 20 are used on a day-to-day basis, 5 are spares. The district wants to be fairly certain that no more than 5 buses will ever be tied up in the repair shop, so that bus routes can always be covered. Records show that all 25 buses need repairs on the average of one bus every 10 days. What is the minimum service rate that will be required at the repair shop if the district wants to be 85 percent certain that no more than 5 buses are in the repair shop at one time?

   HINT: The probability that no more than five buses will be in the repair shop is P(0) + P(1) + P(2) + P(3) + P(4) + P(5). (Why?) Hence, we must find the least possible S, accurate to one-tenth (one decimal place).

OVERVIEW OF QUEUEING THEORY

Queueing theory is a useful decision making tool for the administrator. There are many waiting line situations in education—students seeing a counselor, teachers using a film projector, buses being repaired, supply requests being processed, patrons phoning the school office, and so forth; thus there are many situations that lend themselves to analysis through queueing theory. The technique can be used to analyze an existing waiting line situation (for example, teachers complain that they have to wait too long to use terminals for computer-assisted instruction and the administrator must decide how many terminals will be needed to adequately serve the students.)

Most queueing theory formulas are not too difficult to solve, so many queueing problems can be solved by hand. Computer programs, however, greatly speed their solution and are especially convenient when the administrator wishes to analyze a problem in a number of ways using different parameters. As with the other operations research techniques, computer

ADVANTAGES AND LIMITATIONS OF QUEUEING THEORY

Queueing theory is a mathematical tool that can easily be used by administrators, efficiency experts, and other decision makers to analyze waiting-line problems more fully. It can be applied to any situation which satisfies certain basic conditions assumed in the mathematics of the theory. Armed only with the basic parameters of the system, the administrator can calculate several statistics that are useful in cost-utilization studies.

Queueing theory has several distinct benefits for administrators:

1. It provides them with a framework for studying the current queueing system, enabling them to isolate problem areas of the existing system and discover possible solutions.
2. The statistics provided can function as criteria for determining the "best" solution to a problem.
3. The resulting data may suggest possible changes in the system.
4. They are able to test several alternative modifications to the system quickly and at small cost, in order to determine the most economical and/or practical course to pursue.

Some of its limitations are:

1. The problems it can handle must satisfy certain conditions, and even then the results are only approximations to real-world situations.
2. It is a very limited tool whenever small populations (less than 30) are involved.
3. The results it gives are *average* figures: it is impossible to determine, for example, *exactly* how long a customer will have to wait in line.
4. Queueing theory yields data but does not to describe real-world situations. The final limitation is all too familiar to anyone in a management position. One can gather volumes of data on a particular problem to help determine the best course of action; the final decision, however, is made not by mathematics but by humans, who must take ultimate responsibility for any decisions made.

SUMMARY OF THE MAIN POINTS IN QUEUEING THEORY

Parts of a Queueing Theory Problem

1. Data to collect about the problem:
   - Service facility: location at which service is rendered (F).
   - Customer: user of the service facility.
   - Source field: population of potential customers:
     - Finite (less than 30 potential customers—denoted by M);
     - Infinite (30 or more potential customers).
   - Arrival rate: average number of customers arriving at the facility per unit time (A).
   - Service rate: average number of customers serviced by a facility per unit time (S).
2. Statistics to be calculated using QUEUE:
   - P(O): probability that the facility is idle (no customers being serviced and no customers waiting).
   - E(W): expected number of customers waiting for service.
   - E(T): expected waiting time for an arrival.
   - E(N): expected number of customers waiting for service or being serviced.
   - P(N): probability that N users are being serviced or waiting for service.
   - P(N > L): probability that more than L users are being serviced or are waiting.

Conditions a Queueing Theory Problem Must Satisfy

1. Arrivals of customers are random and

5. Customers entering the queueing system do not leave until they are serviced.
6. Rate of arrivals must be less than the rate of service ($A < S$ for single facility systems; $A < F \cdot S$ for multiple facility systems).

es of Queueing Theory Problems

1. Infinite population, single service facility.
2. Infinite population, multiple service facilities;
   Additional conditions needed: 1. Each facility must have the same service rate; 2. Only a single waiting line may be formed.
3. Finite population, single service facility.

MMARY OF EQUATIONS FOR EUEING THEORY

ations for Infinite Population, le Service Facility

) $= 1 - \frac{A}{S}$ — Probability that service facility is idle.

V) $= \frac{A^2}{S(S - A)}$ — Expected number of customers waiting for service.

T) $= \frac{A}{S - A}$ — Expected number being serviced or waiting for service.

) $= \frac{E(W)}{A} = \frac{A}{S(S - A)}$ — Expected waiting time for an arrival.

) $= \left(\frac{A}{S}\right)^N P(O)$ — Probability of N customers being serviced or waiting.

Equations for Infinite Population, Multiple Service Facilities

$$P(0) = \left[\frac{(A/S)^F}{F!\left(1 - \frac{(A/S)}{F}\right)} + \sum_{i=0}^{F-1} \frac{(A/S)^i}{i!}\right]^{-1}$$

Probability that all service facilities are idle.

$$E(W) = P(0)\,\frac{(A/S)^{F+1}}{F \cdot F!\left(1 - \frac{(A/S)}{F}\right)^2}$$

Expected number of customers waiting for service

$$E(T) = \frac{E(W)}{A}$$

Expected waiting time for an arrival.

$$P(N) = P(0) \cdot \frac{(A/S)^N}{N!} \quad \text{if } N < F$$

$$P(N) = P(0)\,\frac{(A/S)^N}{F! \cdot F^{(N-F)}} \quad \text{if } N > F$$

Probability that N customers are being serviced or are waiting for service.

Equations for Finite Population, Single Service Facility

$$P(0) = P(M)\frac{1}{M!}\left(\frac{S}{A}\right)^M$$

Probability that service facility is idle.

$$E(W) = M - \frac{(A + S)}{A} \cdot [1 - P(0)]$$

Expected number of customers waiting for service.

$$E(T) = \frac{1}{S} \cdot \left[\frac{M}{1 - P(0)} - \frac{1 + (A/S)}{(A/S)}\right]$$

Expected waiting time for an arrival.

$$P(N) = P(M)\frac{1}{(M-N)!}\left(\frac{S}{A}\right)^{M-N}$$

Probability that N customers (N < M) are being serviced or are waiting for service.

FINAL REVIEW

Exercises

1. Identify those of the following situations that are applicable to queueing theory. For situations that could be analyzed by queueing theory, identify the source field (finite or infinite), the customers, the service facilities, typical arrival and service rates, and what significance the quantities E(T) and E(W) would have (i.e. what would E(T), the waiting time of an arrival, represent in a particular problem?). NOTE: This exercise is meant not only to test your ability to recognize a problem applicable to queueing theory, but also to give you more ideas about where to use queueing theory in educational administration.
   (a) Giving vaccinations to elementary school children in a district.
   (b) School cafeteria lines.
   (c) Scheduling maintenance crews.
   (d) Ticket sales at football games.
   (e) Students waiting to see a counselor.
   (f) Creating a master schedule of classes for a large high school.
   (g) Setting up hearings on the school bond issue.
   (h) Teachers using a projection room to preview films.
   (i) Children waiting for school buses in the morning.
   (j) Deciding whether to hire another secretary for the school district administration.
   (k) Scheduling print-shop orders.
2. Select one of the above situations with which you are familiar, obtain realistic data on arrival and service rates, and use QUEUE to compute expected waiting time and expected number waiting. Based on the results of your initial queueing analysis, does there appear to be a waiting line problem? If so, how would you solve the problem? If the solution involves changing variables, change them and rerun the program to obtain revised queueing figures. Continue your analysis and revision until you have reached an effective decision concerning the solution.

BIBLIOGRAPHY AND SUGGESTED READINGS

Gross, Donald, and Carl M. Harris. *Fundamentals of Queueing Theory.* New York: John Wiley and Sons, 1974.

Harris, Roy D., Michael J. Maggard, and William G. Lesso. *Computer Models in Opera-*

Page, Eric. *Queueing Theory in Operations Research.* London: Butterworths, 1972.
Shamblin, James E., and G. T. Stevens, Jr. *Operations Research: A Fundamental Approach.* McGraw-Hill, 1974. Suggested reading: Chapter 8, Queueing or Waiting Line Theory, pp. 187–224.
Siemens, Nicolai, C. H. Marting, and Frank Greenwood. *Operations Research.* New York: The Free Press, 1973. Suggested reading: Chapter 11, Queueing Theory and Its Applications, pp. 275–326.
Van Dusseldorp, Ralph A., D. E. Richardson, and W. J. Foley. *Educational Decision-Making Through Operations Research.* Boston: Allyn and Bacon, 1971.
Wagner, Harvey M *Principles of Operations Research.* Englewood Cliffs, N.J.: Prentice-Hall, 1975. Suggested reading: Chapter 20, Waiting Line Models, pp. 851–902.

5 COMPUTER SIMULATION

CHAPTER PREVIEW This chapter introduces the technique of simulation as a management tool for decision making. Simulation is a problem-solving technique which allows the user to try out a variety of solutions, on a model of the real-world situation, where using the real-world situation itself would be impractical. It is primarily useful in the frequently encountered situations where experimenting with the real situation is either too expensive or impossible.

Throughout the chapter, practical examples illustrate the application of simulation techniques and the use of computer programs in typical simulations. At key points, exercises, mostly based on sample computer programs, encourage the user to apply the techniques described.

CHAPTER AIMS Successful completion of this chapter with its exercises should provide the reader with a basic understanding of simulation techniques and a good feel for the use of computers in simulations. As a result, the user should be able to readily identify problem situations for which simulation offers an effective approach and be able to use computer simulation programs, where appropriate and available, to aid in solutions.

INTRODUCTION TO SIMULATION

COMPONENTS OF A SIMULATION

Simulation is the science of representing reality by artificial means. The word simulate is defined, "to give the appearance or effect of," so that a technique for simulation can be described as a method for giving the appearance or effect

Suppose you were a school superintendent about to approach the school board with a proposal to outlaw candy and coke machines in favor of vending machines that dispensed healthy food. Your problem is to find the best way to present the argument to the board. One thing you might use is a role-playing simulation would project how they would react to the presentation. In this way, you could find a strategy you feel would work best. By role-playing, you are conducting a simulation.

Other everyday uses of simulation are common. When you try out your ideas on someone else, you are in effect using the other person to simulate some part of the real world for you. A community advisory committee which provides the community reaction to alternative courses of action that the school district is considering is another example of a simulation. The committee is used "to give the appearance or effect of" the reaction of the entire community.

Situations like the above are alike in one respect: *The participants are interested in real situations which are impossible or impractical to examine directly.* The real situation is called the *object system.* The object system of a simulation may be as limited as a single human being or as extensive as the United States economy. In the first example above, it was the school board.

In a simulation, we do not work directly with the object system. Instead, we use a *model* of the object system—a representation of the real situation. The model in the example of approaching the school board regarding healthy food in vending machines was your representation of how the board would react to your various arguments. This model would take into account the personalities and attitudes of each of the board members.

Using a model of the object system allows us to investigate what will probably happen in the real situation under various conditions, called the *input* to the model. In the school board problem, the inputs were your different arguments for healthy food in vending machines. The model of the school board allowed you to predict the board's reaction to each argument. The hypothetical reactions of the school board are the *output* of the simulation. Thus, a model in a simulation can be thought of as a set of rules for determining how inputs are related to outputs.

It should now be evident that the primary purpose of a simulation is to derive information to be used to make inferences about the actual behavior of the object system, without the cost, inconvenience, or danger that might be entailed if the actual object system were used.

In many simulations, several different inputs are tried, to provide the user with a set of alternative outputs. This trial-and-error method locates the output which is most desirable to the user.

Thus we can think of simulation as a circular process involving four stages:

1. Form a question about the behavior of the object system;
2. Express the question in the form of an input to the model;
3. Receive an output from the model;
4. If the output is acceptable, make inferences from it about the object system. If not, go back and try other input.

Another educational application of simulation is the "in-basket" technique used by educational administrators. The object is to train educational administrators in making decisions by having them process solutions to actual school problems. Each trainee has an in-basket. Problems are presented to him by placing written communications in his "in-basket." The trainee must then act on the problem, decide on a course of action, prepare a written statement of the solution, and place it in his out-basket. The statement is reviewed by the leaders of the training session and appropriate feedback is prepared and returned to the trainee through the in-basket. In this case, the model is the representation of a school district by training leaders. Inputs are the administrative trainee's responses to the various questions posed through the in-basket. On the basis of these responses, output or feedback, is provided to the trainee by leaders of the training session. The trainee is supposed to make inferences about the actual operation of a school district from the feedback he receives.[1]

As a final example of simulation, let's consider a different kind of model—one used for prediction rather than training. Of primary im-

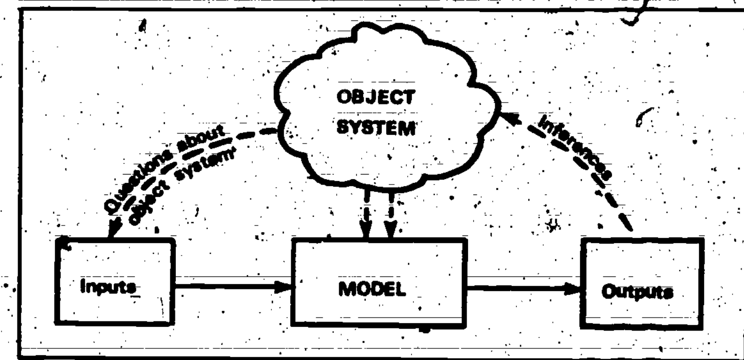


Figure 5-1.
A Generalized Simulation. One example that clearly demonstrates simulation as a process is the typical high school driver education class. Here, simulation is used as a training device. The student sits in front of a mockup of the driver's compartment of a typical automobile. A highway scene is projected in front of the student. The student must "drive" the mockup car so as to stay on the road in the highway scene. The model of the car and highway conditions is provided by the machinery. The input is the driving reactions of the student driver. The output is the feedback to the student by the machine which tells him what is happening to the car. From the feedback, the student infers what would happen if he were driving a real car. Driving is a situation where it would be dangerous for the student to deal initially with the real object system; the simulation allows him to learn safely many of the principles of driving by dealing with only a model of the real situation. The components of the driver training simulation are illustrated in Figure 5-2.

portance to educators is the need to predict future school enrollments. In order to make such predictions, a number of population growth models have been designed. As input they take such information as the expected rate of economic growth, the birth rate, the mortality rate, the distribution of sexes, and the age distribution within the present population. The model is a set of rules (mathematical equations) relating each of these variables to the growth in population for a given geographic area. By using a model of population growth and these input factors, it is possible to output predictions about population changes within the given geographic area. Educators may then make inferences about future needs in education.

It is important to remember, especially with a simulation whose primary purpose is prediction, that output comes only from the model; inferences made about the object system are valid only to the degree that the model accurately reflects the real situation.

The simulation examples given above illustrate the four components of a simulation:

1. We were interested in the behavior of the *object system*, or real situation.
2. The real situation was examined indirectly by using a *model*.
3. Questions were asked about the object system by using them as *inputs* to the model.
4. Inferences were made about the object system by examining the *outputs* of the model.

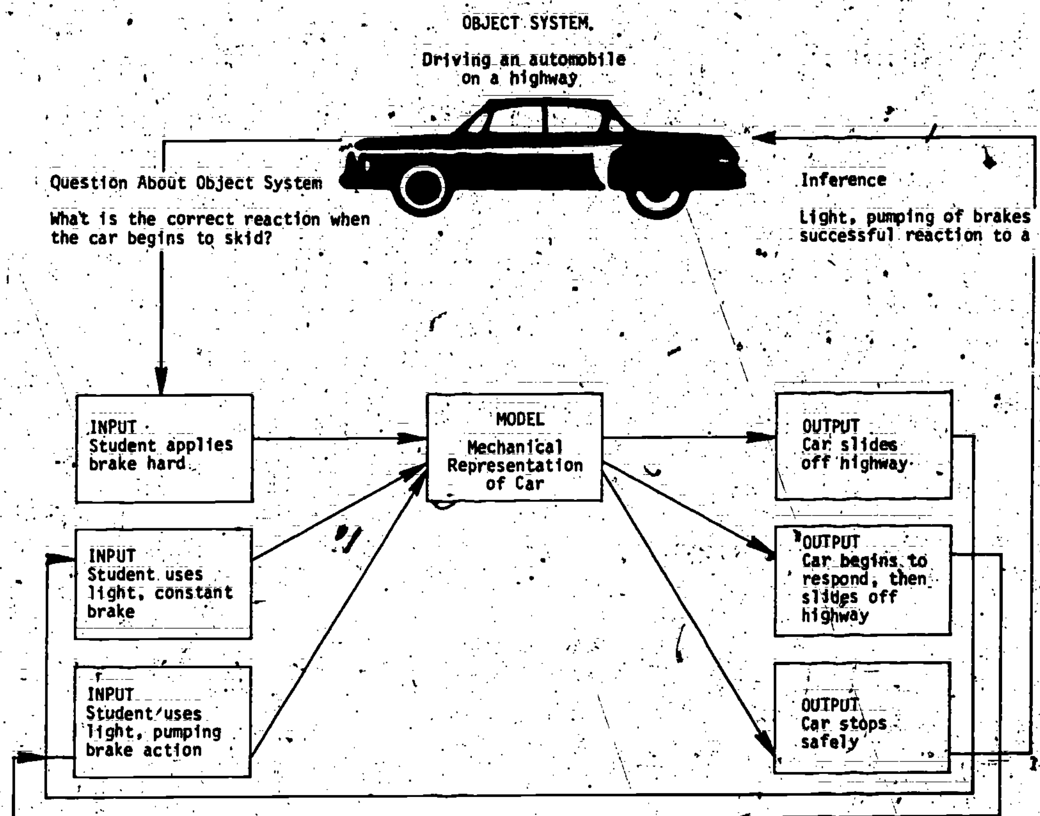


Figure 5-2.
Diagram of a Driver Training Simulation.

REVIEW

Terminology

1. SIMULATION: the exercise or manipulation of a model of the real world.
2. OBJECT SYSTEM: the real situation a model represents.
3. MODEL: a representation of a real situation involving rules for determining how input is related to output.
4. INPUT: conditions applied to a model.
5. OUTPUT: hypothetical reactions of a model to input.

CLASSES AND TYPES OF SIMULATION

Man-model and Man-machine

The examples cited above represent two classes of simulations: *man-model* and *man-machine*; tions are of this class: administrative trainees provide input to a model of a school district by proposing solutions to problems presented to them in their in-baskets; leaders of the simulation take the solutions and, acting as a model of is represented by the driver education example. The model is operated by a machine or computer, but a human interacts with the model by directly providing input and is a vital part of the simulation process.

The use of computers in simulation has become increasingly important for a number of reasons:

1. They can do calculations at great speeds, solving problems in seconds which would take men years to solve by hand.
2. They can process masses of data quickly and can retrieve great volumes of information in seconds.
3. They can solve complex logical problems, acting as an extension of man's own reasoning capacity. The problems, of course, must first be translated into instructions the computer can follow.)

Simulations using the computer thus allow us to run literally thousands of inputs through a model of an object system in relatively short periods of time, enabling us to ask (and answer) a great number of questions about the object system.

Deterministic and Stochastic Simulations

In addition to the division into two classes, simulations may also be classified according to two types—*deterministic* and *stochastic* (sto-kas-tik), or probabalistic. These are based on the processes used in the model to represent the object system. In deterministic simulations, each input and its corresponding output is related by the same set of rules or calculations, so that the same input will always yield the same results, regardless of how many times it is run through the model. The driver education simulation is an example of this type. We know exactly how an automobile (the object system) will react to movements of the steering wheel and control pedals and the model is designed to react the same way. Thus, it is determined what the outputs of the model will be when the student in the simulation moves the steering wheel or steps on the accelerator or brake. When the relation-model will be. Stochastic means random; and stochastic simulations are simulations in which some random element has been incorporated into the model. This means that running the same input through the model a number of times will not necessarily yield the same output every time. For an example, consider the population growth simulation. In this model, population growth is affected by a number of variables, but within certain limits it is random: though population is growing, and is affected by the birth rate and the mortality rate, it is also affected by a number of other factors. We could build a model which assumes that the effects of all other factors besides birth and mortality rates are essentially random. Therefore, a random variable could be used to represent these factors and it would be incorporated into the calculation of every output. The random variable differentiates the stochastic simulation from the deterministic simulation.

In many cases, the designer of a model has a choice between making it a man-model or man-machine and between making it deterministic or stochastic. The decision is influenced by the nature of the object system and the degree of accuracy the model must reflect for the object system. To illustrate this, let's consider the following examples.

Suppose you are trying to decide whether or not to increase the number of buses in your district and, if so, by how many buses. One of the factors in your decision will be the cost to operate each bus for one day. You want to include the increased number of drivers needed, the initial outlay for buying the additional buses, the reduced cost of gasoline a local dealer will give the school district for buying in quantity, and the possible increase in the maintenance crew. You could set up a man-model and ask the foreman of the maintenance crew to estimate the cost per bus per day for various numbers of extra buses (you provide input). Depending on the foreman's experience, his outputs might be accurate enough for you to make your decision; it is possible, however, that you need a more accurate cost figure. In that case you would probably have someone condistrict and each of the four factors listed above, and then, using these relationships, compute the output of cost per bus per day. The model could also be stochastic if, for instance, the maintenance variable had some random allowance for mechanical failures rather than a standard estimate per bus. The decision on which type of model you would choose for your simulation would depend on how closely you want the outputs to approximate the object system.

Some situations (object systems) naturally lend themselves to one type or class of simulation over the other possibilities. A simulation for training school administrators naturally tends toward a man-model (like the in-basket method mentioned earlier) because there are so many human factors involved in each decision. On the other hand, simulations for such situations as scheduling classes, determining cost per unit, and population projections—which are readily quantifiable—tend to call for man-machine models. Some cut-and-dried situations like deciding on where to cut a budget suggest using deterministic models. Others, such as projecting the need for substitute teachers, are almost completely random and demand a stochastic model.

The difference inherent in object systems means that although there are four different combinations of type and class, it is not always possible to construct four different models for the same object system. In most cases, the characteristics of the object system require that some of those combinations be eliminated at the start of the simulation process. Thus, the designer of the model usually has only one or two real choices, and they are governed by what the output of the model is to be used for.

Simulation of an object system is useful for a number of purposes. We can use it to explain past behavior of the object system, to teach about the object system, and to predict future behavior of the object system. In this course we shall be primarily interested in predicting future behavior, because we are using simulation as an aid for decision making.

REVIEW

Terminology

1. MAN-MODEL: class of simulations in which humans are directly involved in both the input to and operation of the model.
2. MAN-MACHINE: class of simulations in which the model is operated by a machine or computer, while humans provide input.
3. DETERMINISTIC SIMULATION: type of simulation in which the relationship between inputs and outputs is completely determined.
4. STOCHASTIC SIMULATION: type of simulation in which a random variable is incorporated into the calculation of every output.

Exercises

1. What are the four stages in the simulation process? (Review text page 124–125.)
2. State three reasons for computers' increasing importance in simulations. (Review page 127.)
3. A school will have six aides assigned to it for the next school year, each of whom is qualified for each of the six positions available. Because of varying seniority, one aide makes \$4/hour, two make \$3.70/hour, and the other three make \$3.25/hour. Two of the positions are full-time (eight hours per day) and the other four are three-fourths time (six hours per day). The principal must include the aides' salaries in his or her esti-

assignments. He or she decides to assign each side a number from 1 to 6 and rolls a die. The first two sides whose numbers are rolled are assigned to the full-time positions. Is this simulation based on a man-model or a man-machine model? Is it stochastic or deterministic? (Review text pages 126–128.)

USING A SIMULATION

ROUTING SCHOOL BUSES USING A SIMULATION

Let's now consider an example of operating a simulation. A matter often faced by educational administrators is how to arrange bus schedules. The problem is usually to determine the best route for a bus to follow in its pickups and deliveries so that all children are picked up and costs of transportation are minimal. Normally an administrator will sit down with a large map of the school district and attempt to plan bus routes by hand. The main criterion will be that all children should be picked up, and such matters as cost will make little difference at the outset. But consider the problem more closely. Out of the hundreds of possible routes each bus can follow, the administrator will most probably pick out only one or two. Also, he or she can plan only one route at a time. Using a computer, however, he or she might be able to try hundreds of different routes to see which is the best.

Components of the Bus Routing Simulation

The first step the administrator must take is to state the problem. In this case, the problem statement is: What is the best way to schedule bus pickups to minimize costs and maximize utility of the buses?

To solve the problem, we can design a simulation. In order to do so, we must first decide what the input variables will be and what output variables we want from the model; then we must determine the rules for relating the inputs to the outputs.

Let's first consider the inputs to the simulation model. In deciding what variables are to be

This information makes up the *constraints*: that is, the information about bus stop locations and numbers of children to be picked up at each stop are constant from one simulation run to the next. Other available input variables are the number of buses used, their capacities, the cost per mile of running the buses, and their average speed. Since the administrator has some control over these inputs, they are called *controllable variables*.

Now consider the outputs the administrator wants from the simulation—what various solutions can be found to the bus routing problem. Obviously, one wants to know the bus routes, i.e. the path each bus will follow and the stops it will make. One probably also wants to know how long it will take each bus to complete its route. Finally, he or she wants to know the cost. Thus, the actual bus routes, the time of the routes, and costs make up the output for the simulation.

In this example, we will assume that the administrator has a deterministic model of bus routing which will relate the inputs to the desired outputs. To be used most efficiently, this model is best represented by a computer program which routes buses by taking the input information and generating the required output information. Such a program exists and is called BUSRTE. The actual model used in the programs—a collection of mathematical equations and techniques—will not concern us here. It is sufficient to know that BUSRTE will take the input information and give, as output, the bus routes and the time and cost of each.

Preparing Input Data for BUSRTE

The input information must first be put into a

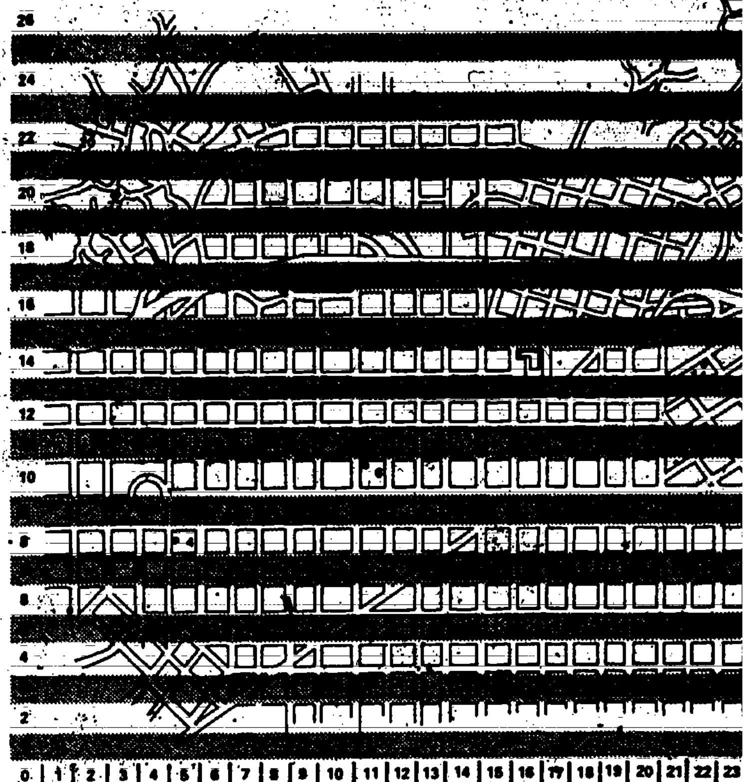


Figure 5.3.
School District Map, Digitized For Use with BUSRTE.

This is done by laying a grid over the map as shown in Figure 5-3. Each column and row of the grid should be numbered as illustrated. To digitize a bus stop, we identify the square it is in. For instance, bus stop number 18 in the upper right-hand corner lies in the square which is in row 25 and column 27; thus its digitized coordinates are 25, 27. A similar process must be carried out with each bus stop until we have the list as given in Figure 5-4. The next step is to record the information for the number of children at each stop. We may know that six children must be picked up at stop number 15; we would record this information by placing a 6 beside the coordinates for stop 15, as shown in the fourth column of Figure 5-4. This process, too, is carried out for each stop on the map.

BUSRTE must have this information provided in a specific form called a DATA state-

The DAT
consecutively
of a correct D

9

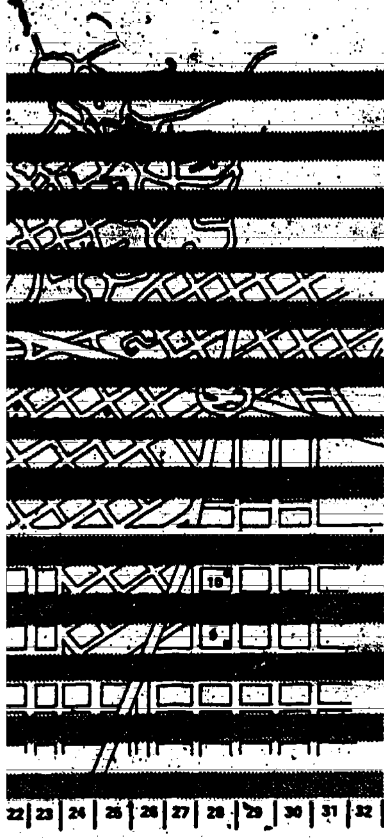
Each stateme
bus stop, so
ments as ther

After the i
been provide
sary:

staten

This stateme
necessary inf
execution ma

Now, let's
our bus routi



ATA statements must be numbered vely, starting with 9000. An example ct DATA statement for BUSRTE is:

9009 DATA 10, 9, 31, 3.

ement contains the information for one so there will be as many DATA statethere are bus stops.

he information for the last bus stop has rided, a final DATA statement is neces-

atement number DATA 0,0,0,0.

ement tells the program that all the information is now present and that may begin.

let's look at all the data involved in outing example. Take the first bus stop

Stop	Vertical Coord.	Horizontal Coord.	No. of Students
1	5	4	5
2	6	8	10
3	7	8	2
4	9	5	4
5	9	12	4
6	12	11	9
7	12	13	10
8	13	18	3
9	6	31	8
10	9	31	3
11	10	28	6
12	12	29	11
13	15	26	15
14	15	23	2
15	20	23	6
16	20	26	15
17	23	26	10
18	25	27	5
19	24	25	2
20	23	2	6
21	22	2	10
22	21	3	8
23	18	9	7

Figure 5–4.
Bus stop coordinates and students at each stop.

The second stop, with coordinates 6 and 8 and ten children, has the following DATA statement:

```
9001 DATA 2,6,8,10
```

The second DATA statement has statement number 9001, because DATA statements are numbered consecutively. Notice that commas and no spaces are placed between each data item just as the DATA statement form shown on page 00 requires.

The complete list of DATA statements for the bus stop map shown in Figure 5–3 is given below.

```
9000 DATA 1,5,4,5
9001 DATA 2,6,8,10
9002 DATA 3,7,8,2
9003 DATA 4,9,5,4
9004 DATA 5,9,12,4
9005 DATA 6,12,11,9
9006 DATA 7,12,13,10
9007 DATA 8,13,18,3
9008 DATA 9,6,31,8
9009 DATA 10,9,31,3
9010 DATA 11,10,28,6
9011 DATA 12,12,29,11
```

```
9017 DATA 18,25,27,5
9018 DATA 19,24,25,2
9019 DATA 20,23,2,6
9020 DATA 21,22,2,10
9021 DATA 22,21,3,8
9022 DATA 23,18,9,7
9023 DATA 0,0,0,0
```

BUSRTE allows us to enter complete sets of such data for any new bus routing problem. For our use in this text, however, BUSRTE also allows the option of using the sample data (shown above) without having to enter each set of data. We will first look at BUSRTE using the sample data above, which are already stored in the program.

Running the BUSRTE Program

To run BUSRTE, you must first get access to the program by typing:

```
GET-BUSRTE
```

on the teletypewriter.[2] Then, tell the program to begin execution by typing:

After BUSRTE begins to run, it will print

```
BUSRTE
```

and it will then print out the question

```
DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?
```

The BUSRTE program already contains the complete set of bus stop data from Figure 5-4. Let's assume that you wish to run the program using that information. In this case, then, you type in "1" directly after the question mark as is shown below:[3]

```
?1
```

You will then be asked another question:

```
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?
```

If you had already run the program (in the same sitting) and provided your own DATA statements, you would type "1." The program would then tell you to call the program and run it again. The entire conversation with the computer thus far would look like this:

```
.GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?1
CALL AND RUN BUSRTE AGAIN

DONE
```

Assume that you have not entered any new data in this sitting and type "2" after the question mark. The program will now continue by requesting from you the coordinates for the location of the school:[4]

```
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?
```

Since the school on our map has a vertical coordinate of 15 (row 15) and a horizontal coordinate of 17 (column 17), you would respond "15,17" after the question mark, as is shown below:

```
?15,17
```

Next, the program will request the capacity of the different types of buses you have available. Let us assume for the sake of this example that you have two types of buses available—48-passenger buses and 36-passenger buses—and you wish to route the 48-passenger buses first. That is, you want the computer to find routes for all of the 48-passenger buses before it routes any of the 36-passenger buses. Examine the printout below to see how the computer requests this information about bus capacities, and how you would respond for this particular example:

```
YOU WILL BE ABLE TO ENTER THE CAPACITY OF EACH VEHICLE
TYPE. WHEN ENTERING, REMEMBER THAT THE COMPUTER WILL
RUN ROUTES IN THE ORDER ENTERED. ENTER ONE AT A TIME
WHEN A QUESTION MARK APPEARS, UP TO 5 ENTRIES. ENTER
0 IF NO MORE ARE DESIRED.
CAPACITY 1
?48
CAPACITY 2
?36
CAPACITY 3
?0
```

Note that you enter "0" to tell the computer that it has all the bus capacities it needs. Note also that you are limited to five bus capacities, and that by reordering the types of buses you enter, you can change the way the buses are routed. In our example since we designated "48" as the first bus capacity and "36" as the second capacity, all the 48-passenger buses will be routed before any of the 36-passenger buses are routed. If we had entered "36" first and "48" second for bus capacities, all the 36-passenger buses would be routed first, followed by the 48-passenger buses.

The computer will next ask you for the number of buses of each capacity that you have specified. Suppose for the sake of the example that you have two 48-passenger buses and four 36-passenger buses. The printout below shows how the computer requests the number of buses and how you would respond in this example:

[3] Underlining is used in the following examples to indicate responses by the user as distinct from computer generated text. No such underlining appears in actual use.

[4] When running BUSRTE, you must make sure that the location of your school is not also a bus stop: the computer will not accept coordinates for the school which are the same as the coordinates for one of the bus stops.

```
ENTER THE NUMBER OF 48   PASSENGER BUSES
?2
ENTER THE NUMBER OF 36   PASSENGER BUSES
?4
```

The computer will next ask for the average speed of the buses. This information is needed to calculate the time taken to complete the bus routes. Let's suppose that the average speed of your buses is 5 miles per hour, taking into account the time necessary to make all the stops and load the students at each stop. The computer's question and the response you would make are shown below:

```
ENTER THE AVERAGE RATE OF TRAVEL IN MPH
?5
```

The computer next needs information about the grid you are using with your map. It must translate the distance between stops on the grid into distance in miles. The computer will ask:

```
ENTER THE NUMBER OF GRID LINES PER MILE.
  FOR EXAMPLE, IF A QUARTER-MILE GRID IS USED, ENTER 4
?12
```

In our example, the distance between grid lines is one block, or 1/12 mile, so we would enter "12" in response, since there are twelve grid lines per mile.

The final data the computer will request is the cost per mile to operate the buses. This may be simply the cost of gasoline, or it may include the bus driver's salary, the cost of maintenance, administrative costs, and insurance. Let's use the figure of $2.25 per mile in our example and assume that all transportation costs—gasoline, maintenance, drivers' salaries, administration, and insurance—are included. The computer's query and your response would appear like this:

```
ENTER COST PER MILE TO OPERATE BUSES
?2.25
```

The computer now has all the information it needs to operate the simulation model and produce the required outputs. Figure 5-5 gives the entire list of inputs exactly as the computer asks for them.

Interpreting the Output of the BUSRTE Program

We now consider the output of this computer simulation. The complete computer output for

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?15,17
YOU WILL BE ABLE TO ENTER THE CAPACITY OF EACH VEHICLE
TYPE.  WHEN ENTERING, REMEMBER THAT THE COMPUTER WILL
RUN ROUTES IN THE ORDER ENTERED.  ENTER ONE AT A TIME
WHEN A QUESTION MARK APPEARS, UP TO 5 ENTRIES.  ENTER
0 IF NO MORE ARE DESIRED.
CAPACITY  1
?48
CAPACITY  2
?36
CAPACITY  3
?0
ENTER THE NUMBER OF 48   PASSENGER BUSES
?2
ENTER THE NUMBER OF 36   PASSENGER BUSES
?4
ENTER THE AVERAGE RATE OF TRAVEL IN MPH
?5
ENTER THE NUMBER OF GRID LINES PER MILE.
  FOR EXAMPLE, IF A QUARTER-MILE GRID IS USED, ENTER 4
?12
ENTER COST PER MILE TO OPERATE BUSES
?2.25
```

Figure 5-5.
Sample Input to BUSRTE Program.

our example is given in Figure 5-6. Let's take the results for one route from Figure 5-6 and cover in detail the output one gets from the computer. The example we will use is Route 1:

```
ROUTE 1
STOP        STUDENTS      MINUTES
 6            9            6.7
 23           7            13
 20           6            21.6
 21           10           22.6
 22           5            24
 4            4            36.2
 S                         49.6
THE COST BASED ON $ 2.25        PER MILE IS $ 9.31
BUS CAPACITY 48    TOTAL STUDENTS 44
```

This output completely describes the route run by the first bus. It details where the bus will stop, in what order the stops will be made, how many children will be picked up at each stop, and how much time will have elapsed from the beginning of the route. This bus stops first at stop number 6, where it picks up nine children, having arrived at the stop 6.7 minutes after leaving the school. The next stop on the route is number 23, where the bus picks up 7 more children. It arrives at stop number 23 thirteen minutes after beginning its route, or 6.3 minutes after leaving the first stop. The bus will proceed until it is filled to capacity or near capacity, then it will head for the school. The last figure under "minutes" gives the time it arrives back at the school; for this route, that is 49.6 minutes after the bus began the route. The cost of this particular bus route is $9.31 and the total number of children picked up is 44.

At the end of the printout in Figure 5-6, some summary information about all bus routes is given. This information includes:

1. *The stops which are not reached by any buses.* In this example, no stops were missed.
2. *The total cost for all routes.* In this example, the cost for all routes was $31.52. This cost will vary depending on the location of the bus stops, the number of students to be bused, and the capacities of the buses.
3. *The total time for all routes.* For this problem, the total time was 168.1 minutes. This figure also will vary with the location of bus stops, the number of students to be bused, the capacities of the buses, and the speed of the buses.
4. *The total number of children picked up and not picked up.* In this example, 161 children were picked up and 0 were not picked up.

```
ROUTE 1
STOP        STUDENTS      MINUTES
 6            9            6.7
 23           7            13
 20           6            21.6
 21           10           22.6
 22           5            24
 4            4            36.2
 S                         49.6
THE COST BASED ON $ 2.25        PER MILE IS $ 9.31
BUS CAPACITY 48    TOTAL STUDENTS 44

ROUTE 2
STOP        STUDENTS      MINUTES
 14           2            6
 11           6            13.1
 9            8            18.1
 10           3            21.1
 12           11           24.7
 13           15           28.9
 S                         37.9
THE COST BASED ON $ 2.25        PER MILE IS $ 7.11
BUS CAPACITY 48    TOTAL STUDENTS 45

ROUTE 3
STOP        STUDENTS      MINUTES
 7            10           5
 3            2            12.1
 1            5            16.5
 2            10           20.7
 5            4            25.7
 8            3            32.9
 S                         35.1
THE COST BASED ON $ 2.25        PER MILE IS $ 6.58
BUS CAPACITY 36    TOTAL STUDENTS 34

ROUTE 4
STOP        STUDENTS      MINUTES
 19           2            12
 18           5            14.3
 17           10           16.5
 16           15           19.5
 S                         29.8
THE COST BASED ON $ 2.25        PER MILE IS $ 5.59
BUS CAPACITY 36    TOTAL STUDENTS 32

ROUTE 5
STOP        STUDENTS      MINUTES
 15           6            7.8
 S                         15.6
THE COST BASED ON $ 2.25        PER MILE IS $ 2.93
BUS CAPACITY 36    TOTAL STUDENTS 6

STOPS WHERE STUDENTS WERE NOT PICKED UP
STOP        STUDENTS
NONE

THE TOTAL COST FOR ALL ROUTES IS $ 31.52
THE TOTAL TIME FOR ALL ROUTES IS 168.1       MINUTES
 161   STUDENTS PICKED UP 0      NOT PICKED UP

WANT TO RUN AGAIN WITH SAME BUS STOP LOCATION DATA?
YES=1,NO=0
?0

DONE
```

Figure 5-6.
Output from BUSRTE.

Notice also at the bottom of Figure 5-6 that when the computer has finished with the problem you are given an option to use the same bus-stop-location data again. If you do not wish to run the program again using this data, type "0" as indicated and the program will end.

Several assumptions made by this computer program should be emphasized:

1. All buses begin and end their routes at the school.
2. Distances used by the program are as the crow flies; that is, the distance between any two bus stops is assumed to be a straight line.
3. The cost of each route is calculated solely on the total number of miles in the bus route and the cost per mile of operating the buses.
4. The time for each route is calculated on the bases of the total number of miles in the route and the estimated speed of the bus.

The BUSRTE simulation is a deterministic, man-machine simulation that takes as input the map of the school bus stops, the number of children at each stop, the number and capacity of buses available, their speed, the cost per mile, and the scale of the grid map. It produces as output a series of bus routes which minimize the distances traveled by each bus and are specified by a list of stops, the total time to run each bus route, and the total cost for each bus route. BUSRTE allows us to experiment with bus routes for a district by varying the capacity and number of school buses available.

Running BUSRTE Using Your Own Bus Stop Data

As we noted above, BUSRTE allows the user the option of entering any new set of bus routing data in order to simulate other bus routing situations, provided you have 24 bus stops or fewer and can assign coordinates to each bus stop location.

If you wish to use BUSRTE in a new situation, you must first organize the new data the way the data in the example were organized. Collect the information in a chart similar to Figure 5-4. With your complete set of data in hand, you are ready to run BUSRTE. Below is an example of how your initial interaction with the program will look.

You should then enter your own DATA statements, using the prescribed format, and type RUN again. When the program begins to run again, it will print out the same first question as it always does:

```
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?
```

You will answer this question by typing 0. The computer will respond with a second question:

```
ARE YOUR DATA STATEMENTS ALREADY ENTERED?
YES =1,  NO =2
?
```

Now you answer by typing 1, since you have

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?0
ARE YOUR DATA STATEMENTS ALREADY ENTERED?
YES =1,  NO =2
?2
FOR EACH BUS STOP, TYPE ONE LINE WITH
THE FOLLOWING FORMAT:
STATEMENT #  DATA  STOP #, VERT COOR, HORIZ COOR, # STUDENTS AT STOP

STATEMENT NUMBERS START WITH 9000 AND ARE
NUMBERED CONSECUTIVELY

EXAMPLE:  9007 DATA 8,14,10,22

AFTER ALL YOUR DATA IS ENTERED, ENTER A
LAST DATA STATEMENT:  90?? DATA 0,0,0,0

24 BUS STOPS ARE ALLOWED

NOW TYPE YOUR DATA STATEMENTS AND RUN THE PROGRAM AGAIN

DONE
```

already entered your new DATA statements. The program then proceeds by asking for the coordinates of the school, bus capacities, and the rest of the information necessary to generate the bus routes. A sample of this conversation appears below:

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?0
ARE YOUR DATA STATEMENTS ALREADY ENTERED?
YES =1, NO =2
?1
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?...
```

The output produced by running BUSRTE with your own data will be in the same format as that shown in Figure 5-6. The actual cost projections, of course, will vary depending on the data you provide.

REVIEW

Terminology

1. CONTROLLABLE VARIABLES: variables in an object system which the administrator or investigator may change or manipulate in order to affect the outcome of the simulation.
2. CONSTRAINTS: relationships among the controllable variables which limit their possible values and which the administrator or investigator is not at liberty to alter without changing the substance of the problem.

Exercises

1. In the example from the text, we determined that it would cost $31.52 and take a total of 168.1 minutes to bus all students in the district to school each day, using two 48-passenger buses and four 36-passenger buses. Now, assume you have only four 48-passenger buses and run the simulation again, using the same bus stop location data. Compare the cost and time figures for the new routes to the ones obtained in the example. Continue to assume that there are 12 grid lines to the mile, that buses travel 5 miles per hour, and that the cost of busing per mile is $2.25.
2. Now, assume you have only five 36-passenger buses. Use BUSRTE to again determine total cost and time figures for bus routes. Then, decide which of the three situations seems most efficient in terms of time and cost: the situation of having four 48-passenger buses; five 36-passenger buses; or two 48-passenger buses and four 36-passenger buses. What other considerations would influence your final decision on the number and type of buses you should purchase?

AN EXPERIMENT USING THE SCHOOL BUS ROUTING SIMULATION

Suppose that a fire completely destroyed the school in the preceding example. The district has the option of rebuilding the school on the same grounds (15, 17) or building a new school on one of two other locations. Since the cost of rebuilding is essentially the same at any location, one of the criteria for site location designated by the school board is that the cost of busing the students should be no higher than it was for the old school and the amount of time the children must ride the bus should be as low as possible.

We can use BUSRTE to provide relevant information for deciding where the new school should be built. All that is necessary is to run BUSRTE changing the location of the school each time while keeping all the other variables constant.

The alternate locations for the school building are given by the coordinates 17, 10 and 13, 16. Suppose we run BUSRTE using these coordinates, assuming two 48-passenger buses and four 36-passenger buses are to be used which have an average speed of 5 mph and a cost per mile of $2.25, and assuming that the map grid lines are spaced 12 to a mile. These runs are shown in Figures 5-7 and 5-8.

The following table summarizes the results of the first run and the two we just made.

School Location	Total Cost	Time of Shortest Route	Time of Longest Route	Total Time	Average Time Per Route*
15,17	$31.52	15.6	49.6	168.1	33.6
17,10	$30.56	11.7	52.2	163.0	32.6
13,16	$32.52	19.8	46.3	173.4	34.7

*Calculated by hand—divide the total time by number of routes.

As can be seen from the table above, all of the proposed locations will cost about the same and will involve about the same busing time to get the students to the school. The 17, 10 location costs slightly less than the others and takes 5-10 minutes less time. From the point of view of one trip, the differences do not appear very significant. When the savings are realized twice a day (once coming to school and once going home) for every school day of the year (approximately 180 days), they add up to $345.60 and 30.6 hours saved on a yearly basis over the second best location (where the school is now). It would be up to the board to decide whether these savings are important enough to be considered in the final location of the school.

The problem above is an example of a situation where a simulation can provide information about a number of alternative courses of action when the physical limitations of the situation make it impossible to manipulate the object system itself. Clearly, the school district could not rebuild the school in each of the proposed locations just to discover how each of them will affect the busing situation. A simulation is the only practical way to obtain this type of information, and the choice of using hand calculations or a computer program will usually depend on the complexity of the model and data involved.

Another Use for the BUSRTE Simulation

In some instances, simulations not only provide sought information but may point up other questions about the object system which had not been raised previously. For example, in the preceding bus-routing problem, it is possible that while the board tries to decide on a new school location using the BUSRTE information, some of the members may realize for the first time that their students are, on the average, riding school buses well over half an hour each way. They might consider this unacceptable and ask the school district to study the problem and propose solutions that would reduce the average one-way trip time to a certain number of minutes.

Let's look for a moment at an example of this sort. Assume that as a result of the previous experiment, the school board finally decided to build the new school at location 17,10. Assume a further decision that, while cost of busing should remain at a minimum, the new bus routes should be designed so that no student will have to ride more than 35 minutes one way. This means that the bus routes to and from the new school location will have to involve less one-way time for some students than those of the original school did; it also means that purchasing additional buses should not be considered unless it is clear that the present six buses (four 36-passenger buses and two 48-passenger buses) cannot be routed in a manner which fits the one-way time constraint.

To solve this problem posed by our hypothetical school board, we can use the BUSRTE printout for school location 17,10 shown in Figure 5-7 as a starting point. Notice that Route 1 violates the new one-way time constraint, since the first two students are picked up (Stop 14) 13.2 minutes after the bus started out but are not dropped off at the school until 52.2 minutes after the bus started out—a ride of 39 minutes. For all the other routes, one-way trips take less than 35 minutes. Notice also that Route 1 picks up the most students (45) of all the routes.

In order to bring the routing into line with the new constraint, we seem to need only to reduce the time of Route 1 without lengthening any other route too much. The most obvious way to accomplish this end may be (continuing to use the BUSRTE simulation) to limit the

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA,
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?17,10
YOU WILL BE ABLE TO ENTER THE CAPACITY OF EACH VEHICLE
TYPE.  WHEN ENTERING, REMEMBER THAT THE COMPUTER WILL
RUN ROUTES IN THE ORDER ENTERED.  ENTER ONE AT A TIME
WHEN A QUESTION MARK APPEARS, UP TO 5 ENTRIES.  ENTER
O IF NO MORE ARE DESIRED.
CAPACITY  1
?48
CAPACITY  2
?36
CAPACITY  3
?0
ENTER THE NUMBER OF 48   PASSENGER BUSES
?2
ENTER THE NUMBER OF 36   PASSENGER BUSES
?4
ENTER THE AVERAGE RATE OF TRAVEL IN MPH
?5
ENTER THE NUMBER OF GRID LINES PER MILE.
   FOR EXAMPLE, IF A QUARTER-MILE GRID IS USED, ENTER 4
?12
ENTER COST PER MILE TO OPERATE BUSES
?2.25

ROUTE 1
STOP        STUDENTS       MINUTES
 14            2             13.2
 11            6             20.2
 9             5             25.2
 10            3             28.2
 12            11            31.8
 13            15            36.1
 5                           52.2
THE COST BASED ON $ 2.25        PER MILE IS $ 9.79
BUS CAPACITY 48    TOTAL STUDENTS 45

ROUTE 2
STOP        STUDENTS       MINUTES
 15            6             13.3
 19            2             17.8
 18            5             20
 17            10            22.3
 16            15            25.3
 8             3             35.9
 5                           44.9
THE COST BASED ON $ 2.25        PER MILE IS $ 8.41
BUS CAPACITY 48    TOTAL STUDENTS 41

ROUTE 3
STOP        STUDENTS       MINUTES
 6             9             5.1
 4             4             11.8
 1             5             15.9
 2             10            20.1
 3             2             21.1
 5             4             25.5
 5                           33.8
THE COST BASED ON $ 2.25        PER MILE IS $ 6.33
BUS CAPACITY 36    TOTAL STUDENTS 34

ROUTE 4
STOP        STUDENTS       MINUTES
 23            7             1.4
 20            6             10
 21            10            11
 22            8             12.4
 5                           20.5
THE COST BASED ON $ 2.25        PER MILE IS $ 3.84
BUS CAPACITY 36    TOTAL STUDENTS 31
```

Figure 5-7.
BUSRTE Run Using School Location of 17, 10.

```
ROUTE 5
STOP        STUDENTS        MINUTES
 7             10             5.8
 S                            11.7
THE COST BASED ON $ 2.25        PER MILE IS $ 2.19
BUS CAPACITY 36    TOTAL STUDENTS 10

STOPS WHERE STUDENTS WERE NOT PICKED UP
STOP        STUDENTS
NONE

THE TOTAL COST FOR ALL ROUTES IS $ 30.56
THE TOTAL TIME FOR ALL ROUTES IS 163  MINUTES
 161  STUDENTS PICKED UP 0    NOT PICKED UP

WANT TO RUN AGAIN WITH SAME BUS STOP LOCATION DATA?
YES=1,NO=0
?0

DONE
```

Figure 5-7 Continued

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?13,16
YOU WILL BE ABLE TO ENTER THE CAPACITY OF EACH VEHICLE
TYPE.  WHEN ENTERING, REMEMBER THAT THE COMPUTER WILL
RUN ROUTES IN THE ORDER ENTERED.  ENTER ONE AT A TIME
WHEN A QUESTION MARK APPEARS, UP TO 5 ENTRIES.  ENTER
0 IF NO MORE ARE DESIRED.
CAPACITY  1
?48
CAPACITY  2
?36
CAPACITY  3
?0
ENTER THE NUMBER OF 48   PASSENGER BUSES
?2
ENTER THE NUMBER OF 36   PASSENGER BUSES
?4
ENTER THE AVERAGE RATE OF TRAVEL IN MPH
?5
ENTER THE NUMBER OF GRID LINES PER MILE.
   FOR EXAMPLE, IF A QUARTER-MILE GRID IS USED, ENTER 4
?12
ENTER COST PER MILE TO OPERATE BUSES
?2.25

ROUTE 1
STOP        STUDENTS        MINUTES
 6             9              5.1
 23            7              11.4
 20            6              20
 21            10             21
 22            8              22.4
 4             4              34.6
 S                            46.3
THE COST BASED ON $ 2.25        PER MILE IS $ 8.68
BUS CAPACITY 48    TOTAL STUDENTS 44

ROUTE 2
STOP        STUDENTS        MINUTES
 14            2              7.3
 11            6              14.4
 9             8              19.4
 10            3              22.4
 12            11             26
 13            15             30.2
 S                            40.4
```

Figure 5-8.
BUSRTE Run Using School Location of 13,16. Continued

```
THE COST BASED ON $ 2.25        PER MILE IS $ 7.57
BUS CAPACITY 48   TOTAL STUDENTS 45

ROUTE 3
STOP          STUDENTS        MINUTES
 19              2              14.2
 18              5              16.4
 17              10             18.7
 16              15             21.7
 S                              33.9
THE COST BASED ON $ 2.25        PER MILE IS $ 6.35
BUS CAPACITY 36   TOTAL STUDENTS 32

ROUTE 4
STOP          STUDENTS        MINUTES
 7               10             3.2
 3               2              10.2
 1               5              1[illegible].7
 2               10             1[illegible].8
 5               4              23.8
 6               3              31
 S                              33
THE COST BASED ON $ 2.25        PER MILE IS $ 6.19
BUS CAPACITY 36   TOTAL STUDENTS 34

ROUTE 5
STOP          STUDENTS        MINUTES
 15              6              9.9
 S                              19.8
THE COST BASED ON $ 2.25        PER MILE IS $ 3.71
BUS CAPACITY 36   TOTAL STUDENTS 6

STOPS WHERE STUDENTS WERE NOT PICKED UP
STOP        STUDENTS
NONE

THE TOTAL COST FOR ALL ROUTES IS $ 32.52
THE TOTAL TIME FOR ALL ROUTES IS 173.4      MINUTES
 161  STUDENTS PICKED UP 0    NOT PICKED UP

WANT TO RUN AGAIN WITH SAME BUS STOP LOCATION DATA?
YES=1,NO=0
?0

DONE
```

Figure 5–8 Continued

number of students that can be picked up by the first bus, decreasing the time taken for Route 1. BUSRTE, we must recognize, is programmed to generate the most mile-economical routes possible; therefore, we can reasonably assume that if a given bus picks up fewer students, it will make fewer stops and travel fewer miles, thereby taking less time to complete a one-way trip.

Using BUSRTE, how can we reduce the number of students the first bus will pick up, since there is no special option provided to do this?[5] We know that BUSRTE is programmed to first fill all buses of the capacity first specified in a run, then fill all those of the capacity specified second, and so forth. In our original run in Figure 5–7, we specified first the 48-passenger capacity and second the 36-passenger capacity; this resulted in a 48-passenger bus being used for the problematic Route 1, which in turn resulted in six stops being made, picking up 45 students and requiring 39 minutes to deliver the first students to the school. As long as we continue to specify first a capacity of 48-passengers, BURSTE will continue to create this excessively long route for the first bus. To reduce the capacity of the bus used for Route 1, therefore, we need to reduce the first capacity we specify. We can do this simply by entering our 36-passenger-capacity buses first and our 48-passenger-capacity ones second.[6]

Figure 5–9 shows the computer run for this experiment. We can see from the printout that when we specify the 36-passenger capacity first,

[5] This is one of the built-in limitations of BUSRTE. All programs have such limitations which must be recognized and worked around in order to make maximum practical use of the simulations.

[6] Notice that if 36-passenger buses still turn out to be too large a capacity for the first bus routed, we can trick one program into reducing Route 1 further by specifying first an even lower capacity bus, even though we may plan to keep and use a 36-passenger bus for the route.

only 28 students are picked up on Route 1, and the first 6 students picked up (Stop 11) have to ride for only 31.3 minutes. The one-way trip time for students on Route 1 is now well within the 35-minute maximum allowed by the board. A look at the rest of the BUSRTE run confirms that this reduction in pick-ups and time for Route 1 does not add excessive time to any other of the routes; all students are picked up and all routes are now within the board's new time constraint.

From this example, you can see that a simulation like BUSRTE can be helpful both in pointing out problematical features involved in existing object systems and in experimenting with solutions for them.

```
GET-BUSRTE
RUN
BUSRTE

DO YOU WISH TO ENTER YOUR OWN BUS STOP DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
ENTER THE SCHOOL COORDINATES, SEPARATED BY A COMMA.
THE VERTICAL COORDINATE SHOULD BE FIRST
?17,10
YOU WILL BE ABLE TO ENTER THE CAPACITY OF EACH VEHICLE
TYPE.  WHEN ENTERING, REMEMBER THAT THE COMPUTER WILL
RUN ROUTES IN THE ORDER ENTERED.  ENTER ONE AT A TIME
WHEN A QUESTION MARK APPEARS. UP TO 5 ENTRIES, ENTER
0 IF NO MORE ARE DESIRED.
CAPACITY  1
?36
CAPACITY  2
?48
CAPACITY  3
?0
ENTER THE NUMBER OF 36   PASSENGER BUSES
?4
ENTER THE NUMBER OF 48   PASSENGER BUSES
?2
ENTER THE AVERAGE RATE OF TRAVEL IN MPH
?5
ENTER THE NUMBER OF GRID LINES PER MILE.
   FOR EXAMPLE, IF A QUARTER-MILE GRID IS USED, ENTER 4
?12
ENTER COST PER MILE TO OPERATE BUSES
?2.25

ROUTE 1
STOP        STUDENTS        MINUTES
 11            6              19.3
 9             8              24.3
 10            3              27.3
 12            11             30.9
 5                            50.6
THE COST BASED ON $ 2.25       PER MILE IS $ 9.48
BUS CAPACITY 36   TOTAL STUDENTS 28

ROUTE 2
STOP        STUDENTS        MINUTES
 19            2              16.6
 18            5              18.8
 17            10             21
 16            15             24
 5                            40.3
THE COST BASED ON $ 2.25       PER MILE IS $ 7.56
BUS CAPACITY 36   TOTAL STUDENTS 32

ROUTE 3
STOP        STUDENTS        MINUTES
 15            6              13.3
 13            15             19.2
 14            2              22.2
 8             3              27.6
 7             10             32.7
 5                            38.5
THE COST BASED ON $ 2.25       PER MILE IS $ 7.22
BUS CAPACITY 36   TOTAL STUDENTS 36
```

Figure 5-9.
BUSRTE for Location 17,10 with 36 Passengers. Continued

```
ROUTE 4
STOP          STUDENTS        MINUTES
 4               4              9.4
 1               5              13.6
 2               10             17.7
 3               2              18.7
 5               4              23.2
 6               9              26.3
 S                              31.4
THE COST BASED ON $ 2.25          PER MILE IS $ 5.89
BUS CAPACITY 36    TOTAL STUDENTS 34

ROUTE 5
STOP          STUDENTS        MINUTES
 20              6              10
 21              10             11
 22              8              12.4
 23              7              19.1
 S                              20.5
THE COST BASED ON $ 2.25          PER MILE IS $ 3.65
BUS CAPACITY 48    TOTAL STUDENTS 31

STOPS WHERE STUDENTS WERE NOT PICKED UP
STOP        STUDENTS
NONE

THE TOTAL COST FOR ALL ROUTES IS $ 34
THE TOTAL TIME FOR ALL ROUTES IS 181.3        MINUTES
 161  STUDENTS PICKED UP 0    NOT PICKED UP

WANT TO RUN AGAIN WITH SAME BUS STOP LOCATION DATA?
YES=1,NO=0
?0

DONE
```

Figure 5-9 Continued

REVIEW

Exercises

1. Figure 5-10 on page 143 provides a map of a large rural school district in a Midwestern state. The district plans to construct a new vocational-technical center and three potential sites have been selected, indicated by letters A, B, and C on the map. Part of the criteria for the final selection of a site is that the cost of busing students to the voc-tech center must be as low as possible. The map indicates the location of bus pick-ups within the district and estimates the number of students who will be at each bus stop. For example, bus stop 1 is Germantown, where one student will be picked up; stop 2 is Pauline, where there will be ten students; and so on. Figure 5-11, on page 144, lists the bus stops, their vertical and horizontal coordinates, and the number of students at each stop.

   Your job is to determine the total costs and times for busing all the students to *each* of the three potential voc-tech center sites. Assume (for all three sites) that (1) you have a fleet of eight mini-buses, which carry 20 passengers each; (2) each bus travels at about 40 miles per hour along its route; (3) the cost of busing will be about 25 cents per mile; and (4) the grid being used has one grid line per mile.

   (a) Which location for the center would require the smallest busing cost and time required for busing?
   (b) In your own opinion, how heavily should these considerations be weighed when the final decision is being made for the location of the voc-tech center?

2. Assume that after reviewing the bus routing information for school location 17,10 provided by BUSRTE, Figure 5–9, the school board decided that it is both possible and desirable to reduce the one-way trip maximum to 25 minutes. Assume that minimum cost continues to be important to the board, so that the purchase of any new buses can be considered only if there is no way to route the present six buses to meet the new one-way time criterion of 25 minutes. Using BUSRTE, plan the best strategy possible to solve this problem.

MAKING ENROLLMENT PROJECTIONS USING A SIMULATION

One important problem the educational administrator faces is planning for the future. They are faced with questions like, how many school buildings will be needed in the next ten years? How must the teaching staff change to meet the changing needs of the community? And how much is the enrollment going to change in the next ten years? All these questions relate to the population changes within the school district. In order to answer them, the administrator must be able to predict changes in population. More precisely, they must be able to predict changes in school enrollment for all grades in their school districts.

Until recently, such predictions were done by intuition or long and laborious hand calculations. But with the advent of computers, new methods have evolved which greatly simplify the

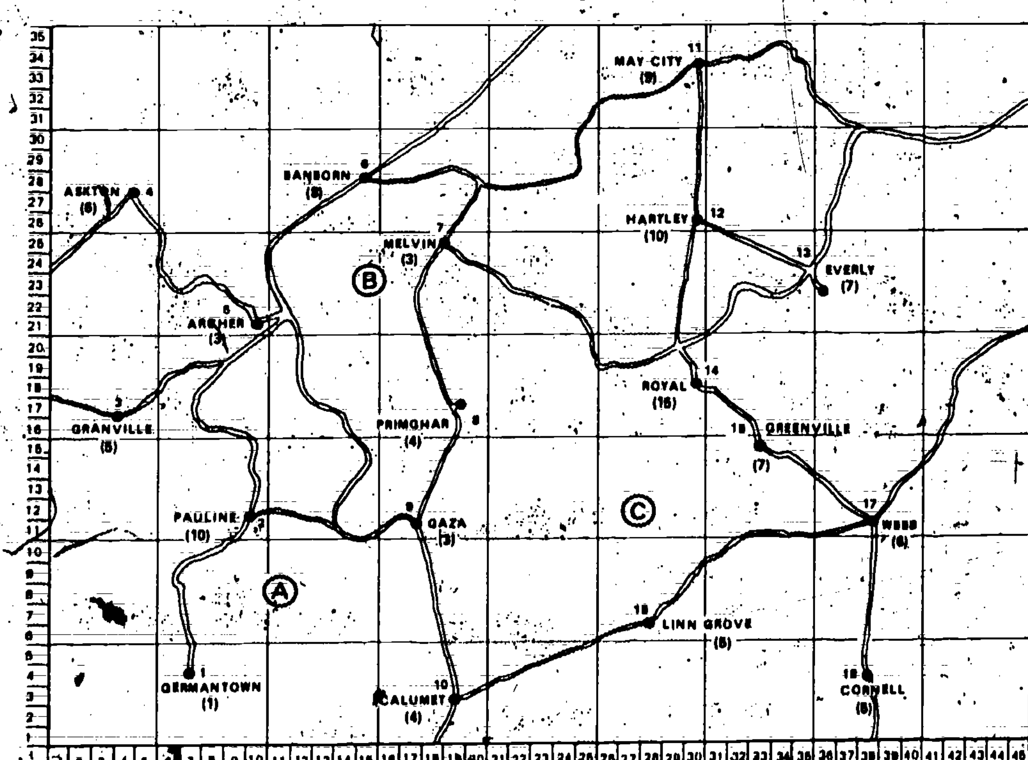


Figure 5–10.
Road Map for Exercise 1, page 142.

	Bus Stop	Vertical Coordinate	Horizontal Coordinate	Number of Students
1	Germantown	4	7	1
2	Pauline	11	9	10
3	Granville	16	3	5
4	Askton	27	4	6
5	Archer	21	10	3
6	Sanborn	28	15	8
7	Melvin	24	19	3
8	Primghar	17	20	4
9	Gaza	11	18	3
10	Calumet	3	19	4
11	May City	33	30	9
12	Hartley	26	30	10
13	Everly	23	36	7
14	Royal	18	30	15
15	Greenville	15	33	7
16	Linn Grove	6	28	5
17	Webb	11	39	6
18	Cornell	4	38	5
	Location of Site A	8	11	
	Location of Site B	23	15	
	Location of Site C	12	28	

Figure 5-11.
Bus Stop Data for Exercise.

process. In the following pages we will discuss an effective computer simulation for predicting school enrollment changes.

ANOTHER COMPUTER SIMULATION—ENROLL

The next simulation program we will be discussing is called ENROLL,[7] a deterministic, man-machine simulation. The model used in ENROLL to predict school enrollments is based on year-to-year percent changes. Data from past years are used to predict future enrollments. Essentially, the enrollment in one grade for one year is compared to the enrollment in the next higher grade for the next year. As the author of the simulation states, "The underlying assumption of the . . . method of projecting enrollments is that historical trends in school district enrollments will continue and will be indicative of future enrollment trends."[8]

To illustrate how the model for enrollment projection works, consider the following enrollment data:

	1975-76	*1976-77*	*1977-78 (Projections)*
Grade 3	780	762	—
Grade 4	—	754	?

(The dashes indicate that we are not interested in those particular figures at the moment.)

During 1975-76 there were 780 students in the third grade. In 1976-77 there were 762 students in the third grade and 754 students in the fourth grade. The question we would like to answer is, How many pupils will probably be in the fourth grade in 1977-78?

To find the answer, we need to calculate the percent change in enrollment from grade 3 to grade 4. From 1975-76 to 1976-77, there was a change of 754 − 780 = −26 students. Since there were 780 students to begin with, this is a net change of −26/780 = −.033 = −3.3%. In other words, there was a change of −3.3% in enrollment between the third and fourth grades.

Now we can use this percent change to predict what will happen in 1977-78. Assuming present trends continue, we can expect that there will also be a change of −3.3% between grades 3 and 4 from 1976-77 to 1977-78. Since there were 762 third graders in 1976-77, in 1977-78 there will be a change of 762 × (−3.3%) = 762 × (−.033) = −25 students. If 25 students are "lost" between grade 3 and grade 4, this means that 762 − 25 = 737 students will be projected for grade 4 during 1977-78.

Examine Figure 5-12, which summarizes the calculations performed above. Remember when you look at percent changes between grades that a *negative* percent change indicates decreasing enrollment, and a *positive* percent change indicates increasing enrollment. Remember also that the projected enrollment based on the above method is an educated guess, based on current trends within a school system. Percent changes will differ from grade to grade within a school system and will also differ among school districts.

[7] ENROLL is a revised version of ENRPRO by Robert D. Nelson.

[8] Robert D. Nelson, "ENRPRO—A Kindergarten–Grade 12 Enrollment Projection Program," p. 1.

	1975-76	1976-77	1977-78 (projections)
Grade 3	780	762	---
Grade 4	---	754	737

(780 → 754: -3.3%; 762 → 737: -3.3%)

Percent change from grade 3 to grade 4 = (Students in grade 4 in 1976-77 − Students in grade 3 in 1975-76) / Students in grade 3 in 1975-76

$= \frac{754-780}{780}$

$= \frac{-26}{780}$

= -.033

= -3.3%

Projected enrollment for grade 4 for 1977-78 = Students in grade 3 in 1976-77 + Projected change in enrollment

= 762 + (-.033 x 762)

= 762 + (-25)

= 737

Figure 5-12.
Summary of Calculations.

REVIEW

Exercise

1. Suppose you had the data given below from the school years 1974–75 and 1975–76.
   (a) What is the percent change from fifth to sixth grade?
   (b) What is the projected sixth grade enrollment for 1976–77?

	1974-75	*1975-76*	*1976-77*
Grade 5	200	220	---
Grade 6	---	150	?

How ENROLL Works

The ENROLL computer simulation makes projections based on five years' population and enrollment data. It uses these data to calculate an *average percent change* in enrollment for each grade, and uses this average percent change as a basis for making projections. Instead of counting each year's percent change equally when calculating the average change, ENROLL uses a *weighted average*.

Let's consider an example using weighted averages. Suppose you are a teacher making out quarterly report card grades. You have only three grades for your students this quarter, based on one major test and two quizzes. Suppose your grades for one student are as follows:

Major test	90
Quiz 1	85
Quiz 2	75

In calculating the student's quarterly grade, you feel you should count the major test more heavily than the quizzes, so you decide to use a weighted average. Specifically, you decide that the major test should count twice as much as either of the quizzes. You would then assign weights to the grades as follows:

	Score	*Weight*
Major test	90	2
Quiz 1	85	1
Quiz 2	75	1

Now, you would just multiply each score by its weight, take the sum of these products and divide by four to obtain the student's quarterly average. You divide by four since the sum of the weights is four (2 + 1 + 1 = 4).

	Score		Weight		Weighted Score
Major test	90	×	2	=	180
Quiz 1	85	×	1	=	85
Quiz 2	75	×	1	=	75
					340

$$\text{Weighted average} = \frac{340}{4} = 85$$

The weighted average gives more emphasis to some scores than to others. In the case of the student above, his good score on the major test helped to cancel out the effect of his mediocre scores on the quizzes. With all three scores being of equal importance, his quarterly average would have been

$$\frac{90 + 85 + 75}{3} = \frac{250}{3} = 83.3$$

In calculating a weighted average for percent changes, ENROLL has a set of standard weights *which give more emphasis to trends in enrollment from recent years than trends from earlier years.* Figure 5-13 gives enrollment figures for grades 1 and 2 for the years 1972-73 to 1976-77. Four percent changes have been calculated, and ENROLL uses them to project the enrollment in grade 2 for 1977-78. Examine the data and follow the calculations to see how this is done. The weights 1, 2, 3, and 4 used in Figure 5-13 are the standard weights of the program and will be used automatically in making projections unless you specify otherwise.

Using ENROLL

Now let's work through an example to see how ENROLL works.

The first information the program needs is at least five years' (and at most 20 years') population and enrollment figures for children of ages 0 to 4 and children in kindergarten to grade 12—18 figures for each school year. For this example, suppose that we have six years' data as given in Figure 5-14. These data must be entered into the computer by DATA statements which have the following form:

```
statement
number      DATA age 0, age 1, age 2, age 3,
                 age 4
statement
number + 1  DATA kind., grade 1, grade 2,
                 grade 3, grade 4, grade 5,
                 grade 6
statement
number + 2  DATA grade 7, grade 8, grade 9,
                 grade 10, grade 11, grade 12
```

The information included in the three DATA statements above correspond to one column (one school year) of the table in Figure 5-13. In the first, the data for one school for ages 0 to 4 are entered; in the next, for kindergarten to grade 6; and so forth.

Because of the way ENROLL is organized internally, DATA statement numbers must begin at 6000 and be numbered consecutively. This means your first DATA statement for each school year will be divisible by 3.

Thus, our first set of DATA statements for the school year 1971-72 should look like this:

```
6000    DATA 686,725,771,774,816
6001    DATA 819,595,650,604,598,607,
             576
6002    DATA 555,569,589,602,540,470
```

All of the appropriate DATA statements generated from Figure 5-13 are already entered in ENROLL. Whenever you wish to use ENROLL with your own data, you will need to prepare it in this format and enter it appropriately into the program at the point the program instructs you to do so. We will demonstrate this later on. For the present, we will not need to prepare the complete set of data from our sample shown in Figure 5-14, because the data have already been correctly organized and entered in ENROLL for use in sample runs.

At this point, then, we are ready to run ENROLL. We begin, as with the other programs, by typing GET-ENROLL and then by typing

	72-73	73-74	74-75	75-76	76-77	77-78	
Grade 1	595 -1.18%	605 -4.30%	572 -0.53%	595 -7.56%	564 -4.16%	---	← weighted average of changes
Grade 2	----	588	579	569	550	540	

Years	Percent Change		Standard Weights		Weighted Percent Change
72-73 to 73-74	-1.18% or -.0118	x	1	=	-.0118
73-74 to 74-75	-4.30% or -.0430	x	2	=	-.0860
74-75 to 75-76	-0.53% or -.0053	x	3	=	-.0159
75-76 to 76-77	-7.56% or -.0756	x	4	=	-.3024
					-.4161

Weighted average percent change for projections $= \frac{-.4161}{1+2+3+4}$

$= \frac{-.4161}{10}$

= -.04161

= -4.16%

Projected enrollment for grade 2 in 1977-78 = Students in grade 1 in 1976-77 + Projected change in enrollment

= 564 + (-.04[illegible] x 564)

= 564 + (-23.68)

= 540.32

= 540 students

Figure 5-13.
Calculations for Percent Changes in Enrollment.

	71-72	72-73	73-74	74-75	75-76	76-77
Age 0	686	649	608	586	562	566
Age 1	725	688	658	641	618	621
Age 2	771	751	708	683	658	638
Age 3	774	750	736	695	702	682
Age 4	816	795	739	731	704	684
Kind.	819	768	763	717	719	699
GR 1	595	605	572	[illegible]	564	544
GR 2	650	588	579	569	550	530
GR 3	604	606	571	590	552	532
GR 4	598	589	615	574	585	565
GR 5	607	585	584	609	578	558
GR 6	576	590	582	595	592	572
GR 7	555	582	618	607	603	587
GR 8	569	550	576	597	613	577
GR 9	589	602	588	642	647	627
GR 10	602	585	632	588	647	627
GR 11	540	611	571	599	590	570
GR 12	470	519	581	554	583	563

Figure 5-14.
Sample Population and Enrollment Data for ENROLL.

RUN,[9] the computer will respond with a serie[illegible]s. In the samples below, user respo[illegible] underlined (although they will not appear that way when you are actually running the program).

The first question:

```
DO YOU WISH TO ENTER YOUR OWN ENROLLMENT DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA=0. USE EXAMPLE DATA =1
?1
```

[9] Make sure you use the correct command for your computer.

Assuming for the moment that we wish to use ENROLL with the data from Figure 5-14, the correct response is "1," as shown.

ENROLL then asks:

```
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
```

In our present example, the answer is "2." Here is a sample of the entire conversation thus far:

```
GET-ENROLL
RUN
ENROLL

DO YOU WISH TO ENTER YOUR OWN ENROLLMENT DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
```

Now, ENROLL will print out the question:

```
FIRST YEAR OF DATA (E.G., IF 1965-66, TYPE 1965) ?
```

The response should be typed in right after the question mark. In our example, the first year is 1971–72, so the correct response is 1971, as shown below:

```
FIRST YEAR OF DATA (E.G., IF 1965-66, TYPE 1965) ?1971
```

Notice that you enter only the first year of the school year. The computer will expect this response to any questions about identifying school years. The program will then ask the last year of data and, in our example, the answer is 1976, since our last year is 1976–77:

```
LAST YEAR OF DATA ?1976
```

The next date you must give is the first year for which you wish enrollment projections to be calculated. As we noted before, the first year for projections must be based on five years' data. In our example, this could be either 1976 or 1977, since we entered data for 1971 through 1976 (six years' data). Assuming we chose 1976, our computer interaction on this question looks like this:

```
FIRST YEAR FOR PROJECTIONS ?1976
```

The computer will next ask you if you wish standard weights used when the average percent change is calculated for each grade, or if you wish to use other weights. For now, let us assume that we want the standard weights, as described on pages 145–146. The computer's question and our response, therefore, would look like the following:

```
IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?2
ENTER FOUR WEIGHTS, SEPARATED BY COMMAS, EARLIEST YEARS
?1,2,0,0
```

We will give an example later of a situation where you would not use standard weights.

Finally, the computer will list three different options for reports it can make, and an option for ending the run at this point; it will then ask which option the user wishes. Assuming we first want the Enrollment Projections Report, we would type in "1."

```
REPORT TO BE PRINTED

    ENROLLMENT PROJECTIONS-1
    COMPARISON-2
    YEAR TO YEAR % CHANGE-3
    END-4
?1
```

Figure 5-15 is the entire computer run as it would appear on the teletypewriter paper, through the point where you choose the report to be printed. Figure 5-16, the Enrollment Projections report, is printed directly after you type 1 in response to the report option. Because of the size of the Enrollment Projections Report, it is printed in two parts, one below the other, on the teletypewriter. We have put the two parts side by side on page 149 so that the report is easier to read.

The Enrollment Projection Report. Let's see what information is given in the Enrollment Projections Report in Figure 5-16. The left-hand portion, entitled PAST CENSUS AND ENROLLMENT DATA, repeats the data that was used as input to the simulation. In this case, data are [illegible] the school years 1971–72 to 1975–76. T[illegible]nd portion, FUTURE ENROLLMENT P[illegible]ECTIONS, gives the enrollment projections for the five school years 1976–77 to 1980–81 (recall that you specified 1976 as the first year for projections). Notice that no projections are made for census data (ages 0 to 4). Thus in the column entitled "1976 to 1977," the simulation predicts that there will be 683 kindergarten students in 1976–77, 554 first

```
GET-ENROLL
RUN
ENROLL

DO YOU WISH TO ENTER YOUR OWN ENROLLMENT DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?1
HAVE YOU ENTERED OTHER DATA SINCE YOU CALLED THIS PROGRAM?
YES =1, NO =2, IF IN DOUBT, ANSWER YES
?2
FIRST YEAR OF DATA (E.G., IF 1965-66, TYPE 1965) ?1971

LAST YEAR OF DATA  ?1976

FIRST YEAR FOR PROJECTIONS  ?1976

IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?1

REPORT TO BE PRINTED

   ENROLLMENT PROJECTIONS-1
   COMPARISON-2
   YEAR TO YEAR % CHANGE-3
   END-4
?1
```

Figure 5-15.
Input and Responses for ENROLL2.

	PAST CENSUS AND ENROLLMENT DATA				
AGE OR GRADE	1971 TO 1972	1972 TO 1973	1973 TO 1974	1974 TO 1975	1975 TO 1976
AGE 0	686	649	608	586	562
AGE 1	725	688	658	641	618
AGE 2	771	751	708	683	658
AGE 3	774	750	736	695	702
AGE 4	816	795	739	731	704
0 - 4	3772	3633	3449	3336	3244
KIND	819	768	763	717	719
GR 1	595	605	572	595	564
GR 2	650	588	579	569	550
GR 3	604	606	571	590	552
1 - 3	1849	1799	1722	1754	1666
GR 4	598	589	615	574	585
GR 5	607	585	584	609	578
GR 6	576	590	582	595	592
4 - 6	1781	1764	1781	1778	1755
K - 6	4449	4331	4266	4249	4140
GR 7	555	582	618	607	603
GR 8	569	550	576	597	613
GR 9	589	602	588	642	647

Figure 5-16.
Enrollment Projections Report from ENROLL. Continued

7 - 9	* 1713	1734	1782	1846	1863	*
GR 10	* 602	585	632	588	647	*
GR 11	* 540	611	571	599	590	*
GR 12	* 470	519	581	554	583	*
10 - 12	* 1612	1715	1784	1741	1820	*
7 - 12	* 3325	3449	3566	3587	3683	*
K - 12	* 7774	7780	7832	7836	7883	*

	*		FUTURE ENROLLMENT PROJECTIONS			*
AGE OR GRADE	* 1976 TO 1977	1977 TO 1978	1978 TO 1979	1979 TO 1980	1980 TO 1981	*
AGE 0	*					
AGE 1	*					
AGE 2	*					
AGE 3	*					
AGE 4	*					
0 - 4	*					
KIND	* 683	683	638	618	585	*
GR 1	* 554	526	526	492	476	*
GR 2	* 540	530	504	504	471	*
GR 3	* 539	529	520	494	494	*
1 - 3	* 1633	1585	1550	1490	1441	*
GR 4	* 551	538	528	519	493	*
GR 5	* 582	548	535	525	516	*
GR 6	* 572	576	542	530	520	*
4 - 6	* 1705	1662	1605	1574	1529	*
K - 6	* 4021	3930	3793	3682	3555	*
GR 7	* 609	588	592	557	545	*
GR 8	* 597	603	582	586	551	*
GR 9	* 666	649	655	632	637	*
7 - 9	* 1872	1840	1829	1775	1733	*
GR 10	* 655	674	657	663	639	*
GR 11	* 635	643	662	645	651	*
GR 12	* 570	613	621	639	623	*
10 - 12	* 1860	1930	1940	1947	1913	*
7 - 12	* 3732	3770	3769	3722	3646	*
K - 12	* 7753	7700	7562	7404	7201	*

Figure 5-16 Continued

graders, 540 second graders, 539 third graders, and so on. Similar predictions are given for the school years through 1980-81. Also included at several points are summaries of different grades and age levels: ages zero to four are summarized by the line labeled "0-4," and Grades 1-3, 4-6, kindergarten-6, 7-9, 10-12, 7-12, and kindergarten-12 are also summarized.

This, then, is the Enrollment Projections Report. It gives estimates of enrollments in future years, based upon past trends.

REVIEW

Exercise

1. For practice in reading the Enrollment Projections Report, answer the following questions, referring to Figure 5-16.
   (a) What is the projected enrollment for the sixth grade in 1978-79?
   (b) What is the total projected high school enrollment (grades 10 to 12) for 1979-80?
   (c) What is the general projected trend in enrollment for grade 5 over the next five years?
   (d) How does the enrollment in kindergarten in 1971-72 compare to the projected enrollment in kindergarten in 1980-81?

The Comparison Report

Immediately after the Enrollment Projections Report has been printed, the computer will ask you to choose another report. Let's now select number 2, the Comparison Report. This can be printed only when you have more than five years

```
REPORT TO BE PRINTED

   ENROLLMENT PROJECTIONS-1
   COMPARISON-2
   YEAR TO YEAR % CHANGE-3
   END-4
?2

YEAR TO COMPARE PROJECTIONS WITH ENROLLMENT ?1976

         ACTUAL VS PROJECTED ENROLLMENTS FOR 1976
         BASED ON DATA FROM  1971   THROUGH 1975
```

	ACTUAL ENROLL	PROJECTED ENROLL	ACTUAL - PROJECTED	% ERROR (A-P)/A
KIND	* 699	683	16	2.29
GR 1	* 544	554	-10	-1.84
GR 2	* 530	540	-10	-1.89
GR 3	* 532	539	-7	-1.32
1 - 3	* 1606	1633	-27	-1.68
GR 4	* 565	551	14	2.48
GR 5	* 558	582	-24	-4.3
GR 6	* 572	572	0	0
4 - 6	* 1695	1705	-10	-.59
K - 6	* 4000	4021	-21	-.53
GR 7	* 587	609	-22	-3.75
GR 8	* 577	597	-20	-3.47
GR 9	* 627	666	-39	-6.22
7 - 9	* 1791	1872	-81	-4.52
GR 10	* 627	655	-28	-4.47
GR 11	* 570	635	-65	-11.4
GR 12	* 563	570	-7	-1.24
10 - 12	* 1760	1860	-100	-5.68
7 - 12	* 3551	3732	-181	-5.1
K - 12	* 7551	7753	-202	-2.66

Figure 5-17.
Comparison Report from ENROLL. Continued

of enrollment data, as in our example; if you have more than five, the years after the first five can be overlapped with projected enrollment years, and the actual and projected enrollments can be compared. *The purpose of this comparison is to determine the accuracy of our predictions.* Figure 5-17 is the Comparison Report for our data, for 1976. Notice that you must tell the computer for which year you wish to compare actual with predicted figures. In our example, 1976 is the *only* year for which we have actual *and* projected data.

The Comparison Report is made up of five columns: the row labels, the actual enrollment, the projected enrollment, the actual minus the projected enrollment, and the percent error in the prediction. The ACTUAL ENROLLment column, which shows the observed enrollment for each grade in the comparison year, contains the same data entered in DATA statements for 1976 at the beginning of the simulation. The PROJECTED ENROLLment column contains the simulation's estimates of the enrollment of each grade level in the comparison year; notice that it is the same as the column for 1976–77 in the Enrollment Projections Report (Figure 5-16). The column headed ACTUAL–PROJECTED contains the difference between the actual enrollment and the projected enrollment for each grade. Finally, the column headed % ERROR contains the percentage of error in the projection for each grade and indicates the accuracy of the simulation's projections. From the last column in Figure 5-17 it is evident that most of the projections are off by not more than 5 percent, which is quite good for prediction purposes. The projection for grade 11, however, is off by more than 10 percent; this sizable margin suggests that some important external factor has interrupted the normal pattern of enrollment change. Thus, the Comparison Report can be used to investigate and evaluate the accuracy of the simulation's predictions.

REVIEW

Exercise

1. For practice in reading the Comparison Report, answer the following questions (referring to Figure 5-17):
   (a) Which projection for 1976 was exactly right? (0% error)
   (b) How does the projection for total district enrollment (K-12) compare to the actual enrollment?
   (c) For grade 7, how far off was the projection from the actual enrollment?

The Year-to-Year Percent Change Report

The third report which can be requested is the Year-to-Year Percent Change Report. This is obtained by typing "3" after the REPORT TO BE PRINTED computer inquiry. Shown in its entirety in Figure 5-18, it is a detailed picture of the enrollment and population changes that have taken place over the five years of input data plus one projected year. Let's look at the portion for grades 1 and 2:

```
         PAST CENSUS AND ENROLLMENT DATA PLUS YR TO YR % CHANGES
                    (ALL DECIMAL FIGURES ARE PERCENTS)

  *                                                          *  PROJECTIONS
  * 1971      1972       1973       1974       1975         *     1976
  *  TO        TO         TO         TO         TO          *      TO
  * 1972      1973       1974       1975       1976         *     1977
---------------------------------------------------------------------------
  *                                                          *
1 * 595       605        572        595        564          *      554
  *                                                          *
  *      1.68      -5.46       4.02       -5.21              * -1.77
  *     -1.18      -4.3        -.53       -7.56              * -4.16
  *                                                          *
2 * 650       588        579        569        550          *      540
```

REPORT TO BE PRINTED

ENROLLMENT PROJECTIONS-1
COMPARISON-2
YEAR TO YEAR % CHANGE-3
END-4

?3

PAST CENSUS AND ENROLLMENT DATA PLUS YR TO YR % CHANGES
(ALL DECIMAL FIGURES ARE PERCENTS)

	1971 TO 1972	1972 TO 1973	1973 TO 1974	1974 TO 1975	1975 TO 1976	PROJECTIONS 1976 TO 1977
A0	686	649	608	586	568	
	-5.39	-6.32	-3.62	-4.1		
	.29	1.39	5.43	5.46		4.12
A1	725	688	658	641	618	585
	-5.1	-4.36	-2.58	-3.59		-5.34
	3.59	2.91	3.8	2.65		3.14
A2	771	751	708	683	658	637
	-2.6	-5.73	-3.53	-3.66		-3.19
	-2.72	-2	-1.84	2.78		-.11
A3	774	750	736	695	702	657
	-3.1	-1.87	-5.57	1.01		-6.41
	2.71	-1.47	-.68	1.29		.89
A4	816	795	739	731	704	704
	-2.57	-7.05	-1.08	-3.69		0
	-5.88	-4.03	-2.98	-1.64		-2.94
K	819	768	763	717	719	683
	-6.23	-.65	-6.03	.28		-5.01
	-26.13	-25.52	-22.02	-21.34		-22.86
1	595	605	572	595	564	554
	1.68	-5.46	4.02	-5.21		-1.77
	-1.18	-4.3	-.53	-7.56		-4.16
2	650	588	579	569	550	540
	-9.54	-1.53	-1.73	-3.34		-1.82
	-6.77	-2.89	1.9	-2.99		-1.88
3	604	606	571	590	552	539
	.33	-5.78	3.33	-6.44		-2.36
	-2.48	1.49	.53	-.85		-.13
4	598	589	615	574	585	551
	-1.51	4.41	-6.67	1.92		-5.81
	-2.17	-.85	-.98	.7		-.4
5	607	585	584	609	578	582
	-3.63	-.17	4.28	-5.09		.69
	-2.8	-.51	1.88	-2.79		-.94
6	576	590	582	595	592	572
	2.43	-1.36	2.23	-.51		-3.38
	1.04	4.75	4.3	1.34		2.88
7	555	582	618	607	603	609
	4.86	6.19	-1.78	-.66		1
	-.9	-1.03	-3.4	.99		-.92
8	569	550	576	597	613	597
	-3.34	4.73	3.65	2.68		-2.61

Figure 5-18.
Year-to Year Percent Change Report from ENROLL. Continued

```
     *    5.8         6.91        11.46       8.38       *  8.75
     *                                                   *
   9 * 589      602         588         648        647   *       666
     *                                                   *
     *    2.21       -2.33        9.18        .78        *  2.94
     *    -.65        4.98        0           .78        *  1.24
     *                                                   *
  10 * 602      585         632         586        647   *       655
     *                                                   *
     *   -2.82        8.03       -6.96       10.03       *  1.24
     *    1.5        -2.39       -5.22        .34        * -1.76
     *                                                   *
  11 * 540      611         571         599        590   *       635
     *                                                   *
     *   13.15       -6.55        4.9        -1.5        *  7.63
     *   -3.89       -4.91       -2.98       -2.67       * -3.33
     *                                                   *
  12 * 470      519         581         554        563   *       570
     *                                                   *
     *   10.43       11.95       -4.65        5.23       * -2.23
```

Figure 5–18 Continued

The first line for grade 1, showing numbers of students, gives the enrollment for each of the years of input data plus the projected enrollment for 1976–77. The next two lines are percentage indices of change from year to year and grade to grade. The top number in the second line is the percent change in enrollment from year to year. For example, the figure 1.68 means that from school year 1971–72 to 1972–73 the first-grade enrollment increased by 1.68%, from 595 students to 605, while between 1972–73 and 1973–74 it decreased 5.46% (–5.46), from 605 to 572. The second line of percentages describes the change from year to year of a single group of students. For example, there were 595 students in first grade in 1971–72 (line no. 1); the following year, 1972–73, there were 588 students enrolled in second grade (line no. 2). This represents a decrease in one class of 1.18% (second line of percentages), or –1.18, between grades 1 and 2.

The Year-to-Year Percent Change Report, then, allows us to follow both the changes in enrollment in a given grade from year to year and the changes in a given class of students as it progresses from grade to grade.

The projected percent change from grade 1 to grade 2 is the *weighted average percent change*, as discussed on pages 145–146. The figure –4.16% is the weighted average of the numbers –1.18%, –4.3%, –.53% and –7.56%. The projected percent change within grade 1, –1.77%, is not a weighted average—that is, in the table on page 00, –1.77% is not the weighted average of the figures 1.68%, –5.46%, 4.02% and –5.21%. The projected percent change within grade 1 was calculated *after* the enrollment projection for grade 1 was made. The figure –1.77% is simply the percent change from 564 students in grade 1 in 1975–76 to 554 projected students in grade 1 in 1976–77.

When you decide to end the simulation, you stop ENROLL simply by entering "4" as the choice of report to be printed. The program will then end, and DONE (or some equivalent word) will be typed out:

```
REPORT TO BE PRINTED
    ENROLLMENT PROJECTIONS-1
    COMPARISON-2
    YEAR TO YEAR % CHANGE-3
    END-4
? 4
DONE
```

REVIEW

Exercise

1. For practice in reading the Year-to-Year Percent Change Report, use the report shown in Figure 5–18 and answer the following questions:
   (a) By what percentage did enrollment in kindergarten decrease from 1971–72 to 1972–73?

(b) What is the projected percent change in kindergarten enrollment from 1975–76 to 1976–77?

(c) How does enrollment in grade 8 in 1973–74 compare to enrollment in grade 9 the following year? What percent change was there?

2. Run ENROLL using the same data as in the example above, except request 1977 as the first year for projections. Produce the Enrollment Projections Report and the Year-to-Year Percent Changes Report (you will not be able to produce the Comparison Report, since you will have no comparison year).

Entering Your Own Data

If you wish to enter your own data rather than using that in Figure 5-14, you respond "0" when the computer asks which type of data should be used. The computer will then ask if the data are already entered. If they are, respond "1" (yes), and the program will progress by asking for the first year of data just as it did when data from the example were used. If not, respond "2" (no), and the computer will give instructions for entering data. A sample follows:

```
GET-ENROLL
RUN
ENROLL

DO YOU WISH TO ENTER YOUR OWN ENROLLMENT DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?0
IS YOUR DATA ALREADY ENTERED?
YES =1, NO =2

YOU MUST ENTER BETWEEN 5 AND 20 YEARS OF DATA.
EACH YEAR REQUIRES 3 DATA STATEMENTS
IN THE FOLLOWING FORMAT:

STATEMENT # DATA AGE 0, AGE 1, AGE 2, AGE 3, AGE 4
STATEMENT # DATA KIND, GR 1, GR 2, GR 3, GR 4, GR 5, GR 6
STATEMENT # DATA GR 7, GR 8, GR 9, GR 10, GR 11, GR 12

DATA STATEMENTS ARE NUMBERED CONSECUTIVELY, STARTING AT 6000

HERE IS AN EXAMPLE OF THE FIRST YEAR'S DATA STATEMENTS.
6001 DATA 667,384,[illegible]1,600
6002 DATA 599,4[illegible],6[illegible]0,705,651,749,683
6003 DATA 740,56[illegible],6[illegible]2,738,585,644

NOW ENTER YOUR OWN DATA STATEMENTS AND RUN THE PROGRAM AGAIN

DONE
```

Varying the Weights

ENROLL offers the user the choice of using the standard weights (1,2,3,4) or entering other weights of his own choosing. If you type in "2" instead of "1" in answer to the computer's question, we will be asked to enter four numbers, as indicated:

```
IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?2
ENTER FOUR WEIGHTS, SEPARATED BY COMMAS, EARLIEST YEARS FIRST
?
```

Recall that the standard weights (1,2,3,4) give more emphasis to recent than to earlier enrollment data. Suppose we want to consider all years of data equally important in our projections. We would then use the weights 1,1,1,1 and would enter them as shown:

```
IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?2
ENTER FOUR WEIGHTS, SEPARATED BY COMMAS, EARLIEST YEARS FIRST
?1,1,1,1
```

Suppose we wished to ignore some of our data. Specifically, suppose that the most recent school year showed atypical trends that we wish to leave out of consideration in making projections. If all other years are equally important, we would enter 1,1,1,0:

```
IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?2
ENTER FOUR WEIGHTS, SEPARATED BY COMMAS, EARLIEST YEARS FI[illegible]
?1,1,1,0
```

Figure 5-19 shows how different weights affect the calculation of percent changes. The data used is the same data for Grades 1 and 2 in the Year-to-Year Percent Change Report (Figure 5-18) on pages 153–154. Notice the variation in projections for Grade 2 for 1977–78 using the different weighting systems.

Using the ENROLL Simulation as a Problem Solving Tool

You now know how the ENROLL enrollment

ENROLLMENT DATA

	71-72	72-73	73-74	74-75	75-76	Projections 76-77
Grade 1	595	605	572	595	564	---
Grade 2	---	588	579	589	550	?

Weights Used	Meaning of These Weights	Calculation of Average Percent Change from Grade 1 to Grade 2	Projected Grade 2 Enrollment for 76-77
1,2,3,4 (standard)	Data from recent years more important than earlier years.	$\frac{1(-.0118) + 2(-.0430) + 3(-.0053) + 4(-.0756)}{1+2+3+4} = -4.16\%$	540
1,1,1,1	Data from all years equally important	$\frac{1(-.0118) + 1(-.0430) + 1(-.0053) + 1(-.0756)}{1+1+1+1} = -3.39\%$	545
1,1,1,0	Most recent year ignored; remaining years of equal importance	$\frac{1(-.0118) + 1(-.0430) + 1(-.0053) + 0(-.0756)}{1+1+1+0} = -2\%$	553
0,0,0,1	All years except the most recent ignored	$\frac{0(-.0118) + 0(-.0430) + 0(-.0053) + 1(-.0756)}{0+0+0+1} = -7.56\%$	521
1,1,2,2	Data from the two most recent years twice as important as the two earliest years	$\frac{1(-.0118) + 1(-.0430) + 2(-.0053) + 2(-.0756)}{1+1+2+2} = -3.61\%$	544

Figure 5-19.
Effects of Using Different Weights in ENROLL.

projections simulation operates. In this section, you will learn about some of the problems this simulation can help you solve.

Prediction simulation is particularly useful in determining the future needs of school districts. It can help you predict the future needs for teachers, classrooms, supplies, taxes, and the district budget. Let us consider the problem of determining future staff requirements.

We need to know how many teachers will be required each year for the next five years. To begin with, we can consult the Enrollment Projections Report (Figure 5-16, pages 149-150), showing the enrollment projections for the next five years. From there, it is possible to project staff requirements by dividing the projected enrollment for each grade by the average class size.

The table below illustrates the calculations involved, assuming an average kindergarten class size of 25 students, dividing each year's projected enrollment by this figure, and rounding to the next highest number. (The projected enrollment figures below are taken from the Enrollment Projection Report.) If our predictions are accurate, there will be a slow decline in the number of kindergarten teachers needed over the next five years. We should remember, however, that these figures are only approximate; they are estimates rather than accurate counts. We can get some idea of how much they may be in error by consulting the Comparison Report (Figure 5-17, page 151). The actual kindergarten enrollment for 1976-77 was 699 and the projected enrollment was 683, so that the projection was in error by 2.29 percent, or 16 students. An error of this size would make the error in the teacher estimate off by at most one teacher—probably an acceptable margin of error for such predictions.

Here is another example of the value of ENROLL as a problem-solving simulation. Suppose you have enrollment and population data entered for the last five years (1972-73 to 1976-77) shown in Figure 5-20. Suppose that during the last two years a large Ford plant in your district was phased out and most of the workers' families were relocated in another state. The dramatic decreases in enrollment shown in Figure 5-20 for the last two years are therefore artificial in the sense of being influenced largely

	Kindergarten				
	76-77	*77-78*	*78-79*	*79-80*	*80-81*
Projected Enrollment	683	683	638	618	585
Teachers Needed	28	28	26	25	24

Age or Grade	72-73	73-74	74-75	75-76	76-77
Age 0	671	654	639	551	502
Age 1	598	576	561	507	463
Age 2	720	699	685	625	580
Age 3	490	476	451	411	367
Age 4	551	530	515	489	420
Kindergarten	588	565	544	502	465
Grade 1	615	610	603	562	522
Grade 2	679	661	651	613	573
Grade 3	580	569	560	517	475
Grade 4	693	679	670	620	575
Grade 5	583	576	560	514	479
Grade 6	705	688	668	620	585
Grade 7	675	661	645	600	562
[illegible]	633	620	603	568	521
[illegible]	648	633	628	581	544
[illegible]	595	582	570	531	489
[illegible]	[illegible]	620	604	566	520
[illegible]	[illegible]	666	649	604	571

[illegible] data for plant closure problem.

from an outside source rather than mainly by the standard birth and transitory rates of the district. Now that the plant is closed, assume you can expect enrollment changes to more or less resume in the pattern established before the plant closed. We can use ENROLL to make the needed projections for the coming years (beginning in 1977–78) by adjusting the weights to allow for the added influence of the plant closing.

Using the data from Figure 5–20, the following DATA statements will be entered at the appropriate time into ENROLL:

```
6000 DATA 671,598,720,490,551
6001 DATA 588,615,679,580,693,583,705
6002 DATA 675,633,648,595,638,671
6003 DATA 654,576,699,476,530
6004 DATA 565,610,661,569,679,576,688
6005 DATA 661,620,633,582,620,666
6006 DATA 639,561,685,451,515
6007 DATA 544,603,651,560,670,560,668
6008 DATA 645,603,628,570,604,649
6009 DATA 551,507,625,411,489
6010 DATA 502,568,613,517,620,514,620
6011 DATA 600,568,581,531,566,604
6012 DATA 502,463,580,367,420
6013 DATA 465,522,573,475,575,479,585
6014 DATA 562,521,544,489,520,571
```

Now it is necessary to decide what the new weights will be. Clearly, the change in the last two years should be weighted less heavily than the change in the first three. Several plausible alternatives are open to us. We could choose to ignore the last two years completely and specify weights of 1,2,0,0; or we could count the last two years minimally and specify 4,3,2,1 or 3,4,1,1. For this example, let's choose 1,2,0,0.

We are now ready to run ENROLL using the new data and weights. Figure 5–21 shows the input to the program and Figure 5–22 shows the resulting Enrollment Projection Report.

These examples are only a few of the many possible uses of the simulation ENROLL.

To summarize, ENROLL is an enrollment projection program which requires as input census and enrollment data for at least five previous years. The output is enrollment projections, based on percent changes between grades.

The prediction method used in ENROLL assumes that past trends in enrollment in a district will continue in the future. Changes can be made in the simulation to make this assumption less rigid—for example, the weights can be varied to allow for unusual circumstances. As the author of the original version cautions, however, school districts experiencing marked changes in their communities need to perform a much deeper analysis of enrollment trends. Areas with a great deal of open land and rapid development of various types of dwelling units, for example, could utilize a dwelling-yield type of enrollment projection study. School districts with neighborhoods experiencing abrupt changes in social and/or economic characteristics should study changes in birth rates and family size.

```
GET-ENROLL
RUN
ENROLL

DO YOU WISH TO ENTER YOUR OWN ENROLLMENT DATA
OR USE THE DATA FROM THE EXAMPLE?
ENTER OWN DATA =0, USE EXAMPLE DATA =1
?0
IS YOUR DATA ALREADY ENTERED?
YES =1, NO =2
?2
YOU MUST ENTER BETWEEN 5 AND 20 YEARS OF DATA.
EACH YEAR REQUIRES 3 DATA STATEMENTS
IN THE FOLLOWING FORMAT:

STATEMENT # DATA AGE 0, AGE 1, AGE 2, AGE 3, AGE 4
STATEMENT # OATA KINO, GR 1, GR 2, GR 3, GR 4, GR 5, GR 6
STATEMENT # DATA GR 7, GR 8, GR 9, GR 10, GR 11, GR 12

DATA STATEMENTS ARE NUMBEREO CONSECUTIVELY, STARTING AT 6000

HERE IS AN EXAMPLE OF THE FIRST YEAR'S OATA STATEMENTS.
6000 OATA 667,384,591,600
6001 OATA 599,484,623,705,651,749,683
6002 DATA 740,569,602,738,585,644

NOW ENTER YOUR OWN DATA STATEMENTS ANO RUN THE PROGRAM AGAIN

DONE

6000 DATA 671,598,720,490,551
6001 OATA 588,615,679,580,693,583,705
6002 DATA 675,633,648,595,632,671
6003 DATA 654,576,699,476,530
6004 DATA 565,610,661,569,679,576,688
6005 DATA 661,620,633,582,680,666
6006 OATA 639,561,685,451,515
6007 OATA 544,603,651,560,670,560,668
6008 DATA 645,603,628,570,604,649
6009 DATA 551,507,625,411,489
6010 DATA 502,562,613,517,620,514,620
6011 OATA 600,568,581,531,566,604
6012 OATA 502,463,580,367,480
6013 OATA 465,522,573,475,575,479,585
6014 DATA 562,521,544,489,520,571

RUN
ENROLL

DO YOU WISH TO ENTER YOUR OWN ENROLLMENT OATA
OR USE THE OATA FROM THE EXAMPLE?
ENTER OWN OATA =0, USE EXAMPLE DATA =1
?0
IS YOUR DATA ALREADY ENTERED?
YES =1, NO =2
?1
FIRST YEAR OF OATA (E.G., IF 1965-66, TYPE 1965) ?1972

LAST YEAR OF DATA ?1976

FIRST YEAR FOR PROJECTIONS ?1977

IF YOU WISH TO USE STANDARD WEIGHTS (1,2,3,4) TO
CALCULATE AVERAGE PERCENT CHANGE, TYPE 1
TO ENTER OTHER WEIGHTS, TYPE 2
?2
ENTER FOUR WEIGHTS, SEPARATED BY COMMAS, EARLIEST YEARS FIRST
?1,2,0,0

REPORT TO BE PRINTED

   ENROLLMENT PROJECTIONS-1
   COMPARISON-2
   YEAR TO YEAR % CHANGE-3
   END-4
?1
```

Figure 5-21.
ENROLL Run for Plant Closing Problem.

AGE OR GRADE	1972 TO 1973	1973 TO 1974	1974 TO 1975	1975 TO 1976	1976 TO 1977
PAST CENSUS AND ENROLLMENT DATA					
AGE 0	671	654	639	551	502
AGE 1	598	576	561	507	463
AGE 2	720	699	685	625	580
AGE 3	490	476	451	411	367
AGE 4	551	530	515	489	420
0 - 4	3030	2935	2851	2583	2332
KIND	588	565	544	502	465
GR 1	615	610	603	562	522
GR 2	679	661	651	613	573
GR 3	580	569	560	517	475
1 - 3	1874	1840	1814	1692	1570
GR 4	693	679	670	680	575
GR 5	583	576	560	514	479
GR 6	705	688	668	620	585
4 - 6	1981	1943	1898	1754	1639
K - 6	4443	4348	4256	3948	3674
GR 7	675	661	645	600	562
GR 8	633	620	603	568	521
GR 9	648	633	628	581	544
7 - 9	1956	1914	1876	1749	1627
GR 10	595	582	570	531	489
GR 11	632	620	604	566	520
GR 12	671	666	649	604	571
10 - 12	1898	1868	1823	1701	1580
7 - 12	3854	3782	3699	3450	3207
K - 12	8297	8130	7955	7398	6881

AGE OR GRADE	1977 TO 1978	1978 TO 1979	1979 TO 1980	1980 TO 1981	1981 TO 1982
FUTURE ENROLLMENT PROJECTIONS					
AGE 0					
AGE 1					
AGE 2					
AGE 3					
AGE 4					
0 - 4					
KIND	430	407	417	394	366
GR 1	491	454	430	440	416
GR 2	558	525	485	459	470

Figure 5-22.
Enrollment Projection Report for Plant Closure Problem. Continued

GR 3	* 463	471	443	409	387	*
1 - 3	* 1532	1450	1358	1308	1273	*
GR 4	* 558	567	553	520	480	*
GR 5	* 475	461	468	457	429	*
GR 6	* 558	554	537	545	533	*
4 - 6	* 1591	1582	1558	1522	1442	*
K - 6	* 3553	3439	3333	3284	3081	*
GR 7	* 548	523	519	503	510	*
GR 8	* 513	501	478	474	459	*
GR 9	* 525	517	505	482	478	*
7 - 9	* 1586	1541	1502	1459	1447	*
GR 10	* 489	472	465	454	433	*
GR 11	* 508	508	490	[illegible]	471	*
GR 12	* 545	532	532	[illegible]	506	*
10 - 12	* 1542	1518	1487	1451	1410	*
7 - 12	* 3128	3053	2989	2910	2857	*
K - 12	* 6681	6492	6322	6134	5938	*

Figure 5–22 Continued

Caution is the word when accepting enrollment projections. A "good" enrollment projection depends on the input data and on whether the projection method truly fits a particular school district. A method that seems to give good projections for a few years could become inappropriate as conditions in the district change.[10]

REVIEW

Exercises

1. Rerun the plant closure problem (Produce an Enrollment Projection Report) using weights of 4,3,2,1 and 3,4,1,1. How do the results differ from each other? Which do you think is the "best" set of weights? What is the rationale for your choice? (Note: It is not necessary to enter the new DATA statements every time you want to rerun ENROLL with different weights. After they have been entered once, just respond "1" (yes) when the program asks if the DATA statements are already entered. Do not retype GET-ENROLL between runs or the Data Statements will be lost.)
2. Choose a combination of weights that you feel is "best" and rerun the plant closure problem. Are your results as good as you expected them to be?

AN EXAMPLE OF A STOCHASTIC PROBABILISTIC SIMULATION

The two previous examples of computer simulations presented in this chapter were both deterministic: that is, the input completely determines what the output will be. For example, no matter how many times BUSRTE is run using the same data, exactly the same bus routes

[10] Robert D. Nelson, "ENRPRO—A Kindergarten–Grade 12 Enrollment Projection Program," p. 1.

will be produced as output. When constructing bus routes the model takes nothing into consideration other than the bus stop data, the location of the school, the average speed of the buses, the cost per mile, and the number and capacity of the buses (all of which are supplied by the user of the program and serve as input).

A great many facets of the real world, however, are not deterministic. Very seldom can we state with complete certainty that input A will always guarantee output B. The number of days per year that it rains and the amount of time it takes to get from home to work in the morning are two examples of situations that vary from year to year or from day to day. Normally, we cope with this kind of variance by establishing average figures as ways of estimating certain behaviors: meteorologists establish an average yearly rainfall for a given geographical area; and each of us develops his or her own average estimate of how long it takes to get to work. For many purposes, these averages and estimates are sufficient. But we should not forget that these and many similar situations are basically stochastic or probabilistic in nature, and that knowing the average rainfall will not guarantee, for example, whether or not it will rain on a given day.

Educational administration has its share of probabilistic, or stochastic, situations, including several we have already discussed—queueing situations, enrollment projections, bus routing problems. While deterministic models can often be used, for ease of model-construction and computation, when simulating these situations, others will require models if we are to have realistic and practical results. Let's look at an example of a probabilistic situation and an accompanying computer simulation based on a stochastic model.

Suppose you are studying the availability of substitute teachers for a certain school. You know that the school has 100 teachers and that they are absent an average of 5 percent of the time. You also know that there are 10 substitutes and each is available to work an average of 70 percent of the time. The question you wish to answer is, how much of the time will fewer substitutes be available than are needed?

On first examination, it would seem that there will always be enough substitutes to cover for the absent teachers. Since the teachers are absent an average of 5 percent of the time, we can figure that an average of 5 teachers (5 percent of 100) will be absent on any one day. Since the substitutes are available an average of 70 percent of the time, you can figure that an average of 7 substitutes (70 percent of 10) will be available on any one day. Therefore, each day will have two more substitutes available than are needed and so there will never be a problem of too few substitutes.

The fallacy in this conclusion is that it assumes the situation is deterministic. The average teacher absentee rate being 5 percent does not guarantee that five teachers or fewer will be absent every day. It is possible that 10 teachers will be absent on one day. Similarly, the substitute availability rate of 70 percent does not guarantee that at least 7 substitutes are available every day. The question remains: what are the chances of having a day with less than enough substitutes?

The Computer Simulation SUBST

SUBST is an example of a computer simulation of the stochastic situation posed above. It takes as input the number of teachers, the teacher absentee rate, the number of substitutes, the substitute availability rate, and the number of days to be simulated. It then simulates a day by generating a random number between 0 and 99 for each teacher and each substitute. Since in this example the teacher absence rate is 5 percent, each random teacher number under 5 represents an absent teacher. The computer then counts how many teachers are absent. Since in this example the substitute availability rate is 70 percent, each random substitute number under 70 represents a substitute available to teach on the day being simulated. The computer then counts up the available substitutes. And so on: each day is simulated in the same manner, and the program keeps track of the outcome for each day and sums it up in the form of the table shown in Figure 5-23.

The first column lists the different values for the difference between the number of substitutes needed and the number available. If more are available than are needed (in which case there is no need to worry), the number in the first column is negative. If not enough are available, this number is positive. Consequently, the rows with positive numbers in the first column will be of particular interest to the administrator.

The second column lists how many days each situation happens. For example, in the

# SUBSTS NEEDED LESS # SUBSTS AVAILABLE	# DAYS	% DAYS	CUMULATIVE # DAYS	CUMULATIVE % DAYS
5	1	3.3	1	3.3
4	1	3.3	2	6.6
2	1	3.3	3	10
1	1	3.3	4	13.3
0	4	13.3	8	26.6
-1	4	13.3	12	40
-2	7	23.3	19	63.3
-3	4	13.3	23	76.6
-4	5	16.6	28	93.3
-6	1	3.3	29	96.6
-7	1	3.3	30	100

Figure 5-23.
Sample Output of SUBST.

table in Figure 5-23, there is one day in the simulation when five more substitutes are needed than are available; there are four days when the number available exactly matches the number needed; and there are five days when there are four extra substitutes available. The third column tells what percentage the second-column numbers are of the total number of days simulated.

The fourth and fifth columns list cumulative totals in number of days and percentages of days respectively. The fourth column in Figure 5-23 shows (for the row headed "4" that there is a cumulative total of two days when there will be four or five too few substitutes, and the fifth column shows that this represents 6.6 percent of the days simulated. The bottom entry in the fourth column will always be the number of days simulated (30, in Figure 5-23), and the bottom entry in the fifth column will always be 100 percent.

Running SUBST. In order to use SUBST, you must issue the appropriate commands to get access to and run the program. Once running, SUBST will request the several inputs needed from you. The first two questions, and the responses appropriate for our example, are:

```
HOW MANY TEACHERS?   (MUST BE LESS THAN 500)
?100
WHAT IS THE TEACHER ABSENTEE RATE (%) ?
?5
```

Note that the answer to the second question must be a percentage. The absentee rate in our example was 5 percent, so the correct response is "5." The program will continue with three more questions:

```
HOW MANY SUBSTITUTES?  (MUST BE LESS THAN 200)
?10
WHAT IS THE AVAILABILITY RATE FOR SUBSTITUTES (%) ?
?70
HOW MANY DAYS ARE TO BE SIMULATED?  (MUST BE LESS THAN 190)
?180
```

Be sure your response to the fourth question (substitute availability rate) is also a percentage. At this point, all the necessary information has been supplied to the program.

If the number of teachers and/or the number of days specified is very large, you will receive a message indicating that it may take some time (over five minutes) to run the simulation.[11] If you do not want to wait this long, you have the option of providing different input using a smaller number of teachers or days, so that the simulation will proceed at a faster rate. There is an example of the over-five-minutes message and the option of entering different input:

```
THE RESULTS OF YOUR INPUT WILL PROBABLY TAKE MORE THAN 5 MINUTES
TO BE PROCESSED.  ARE YOU WILLING TO WAIT OR WOULD YOU RATHER
USE SMALLER NUMBERS FOR INPUT?  (WILLING TO WAIT=1, NEW INPUT=2)
?1
```

SUBST will now proceed with the simulation, informing you of its progress as it goes. The final output will be a table similar to the one in Figure 5-23. Figure 5-24 is a sample of an entire run of SUBST.

[11] This type of simulation requires a very large number of calculations—over 60,000 to simulate 100 teachers for one year. Even though computers are very fast, it still takes some time to perform such large numbers of calculations.

```
GET-SUBST

RUN
SUBST

AVAILABILITY OF SUBSTITUTE TEACHERS

HOW MANY TEACHERS?  (MUST BE LESS THAN 500)
?100
WHAT IS THE TEACHER ABSENTEE RATE (%) ?
?5
HOW MANY SUBSTITUTES?  (MUST BE LESS THAN 200)
?10
WHAT IS THE AVAILABILITY RATE FOR SUBSTITUTES (%) ?
?70
HOW MANY DAYS ARE TO BE SIMULATED?  (MUST BE LESS THAN
?180
THE RESULTS OF YOUR INPUT WILL PROBABLY TAKE MORE    MINUTES
TO BE PROCESSED.  ARE YOU WILLING TO WAIT OR WOULD     HER
   LLER NUMBERS FOR INPUT?  (WILLING TO WAIT=1; NEW INPUT=2)

PLEASE WAIT, SIMULATION IS ON THE 10TH DAY
STILL CALCULATING, SIMULATION IS ON THE 20TH DAY
PLEASE BE PATIENT, SIMULATION IS NOW ON THE 30TH DAY
SIMULATION FOR ALL DAYS COMPLETE, CALCULATION OF TABLE VALUES UNDER WAY
DONE WITH CALCULATING CUMULATIVE DAYS
DONE WITH CALCULATING % DAYS
DONE WITH CALCULATING CUMULATIVE % DAYS
```

# SUBSTS NEEDED LESS # SUBSTS AVAILABLE	# DAYS	% DAYS	CUMULATIVE # DAYS	CUMULATIVE % DAYS
6	1	.5	1	.5
5	2	1.1	3	1.6
4	2	1.1	5	2.7
3	3	1.6	8	4.4
2	6	3.3	14	7.7
1	13	7.2	27	15
0	15	8.3	42	23.3
-1	23	12.7	65	36.1
-2	24	13.3	89	49.4
-3	25	13.8	114	63.3
-4	28	15.5	142	78.8
-5	16	8.8	158	87.7
-6	11	6.1	169	93.8
-7	9	5	178	98.8
-8	2	1.1	180	100

```
DO YOU WANT TO RUN THE SIMULATION AGAIN?  (YES=1, NO=2)
?2

DONE
```

Figure 5-24.
Sample Run of SUBST for 180 Days.

REVIEW

Exercises

1. Answer the following questions about Figure 5-24.
   (a) How many days were simulated?
   (b) How many days needed four more substitutes than were available?
   (c) For what percentage of days were two extra substitutes available?
   (d) What was the total number of days for which fewer substitutes were available than were needed?
   (e) What percentage of total days simulated is this figure?

2. Run SUBST two times using the same input as in Figure 5-24.
   (a) How do the tables compare with each other?
   (b) Which of the three is the right table?

A SUBST Example

Suppose you are in charge of finding substitutes for your high school, and the district superintendent wants you to provide assurance that there will be a shortage of substitutes no more than three times during the year. Assuming there are 85 teachers in your school with an absentee rate of 7 percent, and substitutes in your area are available 65 percent of the time, what is the smallest reasonable number of substitutes you must have on call? (Use 180 days = 1 school year.)

This is essentially a trial-and-error problem. Our strategy will be to make educated guesses on the number of substitutes, test the hypothesis by running SUBST, and modify our guess according to the output. In light of the previous example, a logical starting place would be to run SUBST using 10 substitutes. This will give us a reference point. Figure 5-25 is a run of SUBST using this input.

In this try, there were 56 days when there were fewer substitutes available than were needed. This is very far from the goal of three days, so we must run the program again using a larger number of substitutes. Figure 5-26 shows a second run using 20.

We see from Figure 5-26 that if we use 20 substitutes, there is only one day when the number of substitutes needed exceeds the number

```
GET-SUBST
RUN
SUBST

AVAILABILITY OF SUBSTITUTE TEACHERS

HOW MANY TEACHERS?  (MUST BE LESS THAN 500)
?85
WHAT IS THE TEACHER ABSENTEE RATE (%) ?
?7
HOW MANY SUBSTITUTES?  (MUST BE LESS THAN 200)
?10
WHAT IS THE AVAILABILITY RATE FOR SUBSTITUTES (%) ?
?65
HOW MANY DAYS ARE TO BE SIMULATED?  (MUST BE LESS THAN 190)
?180
THE RESULTS OF YOUR INPUT WILL PROBABLY TAKE MORE THAN 5 MINUTES
TO BE PROCESSED.  ARE YOU WILLING TO WAIT OR WOULD YOU RATHER
USE SMALLER NUMBERS FOR INPUT? (WILLING TO WAIT=1, NEW INPUT=2)
?1
PLEASE WAIT, SIMULATION IS ON THE 10TH DAY
STILL CALCULATING, SIMULATION IS ON THE 20TH DAY
PLEASE BE PATIENT, SIMULATION IS NOW ON THE 50TH DAY
SIMULATION FOR ALL DAYS COMPLETE, CALCULATION OF TABLE VALUES UNDER WAY
DONE WITH CALCULATING CUMULATIVE DAYS
DONE WITH CALCULATING % DAYS
DONE WITH CALCULATING CUMULATIVE % DAYS
```

# SUBSTS NEEDED LESS # SUBSTS AVAILABLE	# DAYS	% DAYS	CUMULATIVE # DAYS	CUMULATIVE % DAYS
7	2	1.1	2	1.1
6	1	.5	3	1.6
5	3	1.6	6	3.3
4	5	2.7	11	6.1
3	10	5.5	21	11.6
2	16	8.8	37	20.5
1	19	10.5	56	31.1
0	22	12.2	78	43.3
-1	38	21.1	116	64.4
-2	12	6.6	128	71.1
-3	22	12.2	150	83.3
-4	15	8.3	165	91.6
-5	8	4.4	173	96.1
-6	4	2.2	177	98.3
-7	2	1.1	179	99.4
-8	1	.5	180	100

Figure 5-25.
Sample Run of SUBST Using 10 Substitutes.

```
DO YOU WANT TO RUN THE SIMULATION AGAIN?  (YES=1, NO=2)
?1
HOW MANY TEACHERS?  (MUST BE LESS THAN 500)
?85
WHAT IS THE TEACHER ABSENTEE RATE (%) ?
?7
HOW MANY SUBSTITUTES?  (MUST BE LESS THAN 200)
?20
WHAT IS THE AVAILABILITY RATE FOR SUBSTITUTES (%) ?
?65
HOW MANY DAYS ARE TO BE SIMULATED?  (MUST BE LESS THAN 190)
?180
THE RESULTS OF YOUR INPUT WILL PROBABLY TAKE MORE THAN 5 MINUTES
TO BE PROCESSED.  ARE YOU WILLING TO WAIT OR WOULD YOU RATHER
USE SMALLER NUMBERS FOR INPUT?  (WILLING TO WAIT=1, NEW INPUT=2)
?1
PLEASE WAIT, SIMULATION IS ON THE 10TH DAY
STILL CALCULATING, SIMULATION IS ON THE 20TH DAY
PLEASE BE PATIENT, SIMULATION IS NOW ON THE 50TH DAY
SIMULATION FOR ALL DAYS COMPLETE, CALCULATION OF TABLE VALUES UNDER WAY
DONE WITH CALCULATING CUMULATIVE DAYS
DONE WITH CALCULATING % DAYS
DONE WITH CALCULATING CUMULATIVE % DAYS
```

# SUBSTS NEEDED LESS # SUBSTS AVAILABLE	# DAYS	% DAYS	CUMULATIVE # DAYS	CUMULATIVE % DAYS
2	1	.5	1	.5
1	1	.5	2	1.1
0	4	2.2	6	3.3
-1	4	[illegible]	10	5.5
-2	9	[illegible]	19	10.5
-3	11	6.1	30	16.6
-4	12	6.6	42	23.3
-5	18	10	60	33.3
-6	23	12.7	83	46.1
-7	17	9.4	100	55.5
-8	23	12.7	123	68.3
-9	19	10.5	142	78.8
-10	13	7.2	155	86.1
-11	13	7.2	168	93.3
-12	4	2.2	172	95.5
-13	4	2.2	176	97.7
-14	3	1.6	179	99.4
-15	1	.5	180	100

Figure 5-26.
Sample Run of SUBST Using 20 Substitutes.

available—and that will certainly please the superintendent. But before concluding that we have solved the problem, we must remember we are working with a stochastic model and the same input does not always yield the same results. For this reason we can never guarantee the result. It would be prudent to run SUBST again several times using 20 substitutes and leaving all the rest of the input the same. Figure 5-27 shows the result of one such rerun.

What can we conclude from the last three runs of SUBST? Nothing, absolutely. We have learned that 20 substitutes will *probably* be sufficient to meet the superintendent's criterion. We cannot tell whether fewer than 20 would be likely to suffice.

```
DO YOU WANT TO RUN THE SIMULATION AGAIN?  (YES=1, NO=2)
?1
HOW MANY TEACHERS?  (MUST BE LESS THAN 500)
?85
WHAT IS THE TEACHER ABSENTEE RATE (%) ?
?7
HOW MANY SUBSTITUTES?  (MUST BE LESS THAN 200)
?20
WHAT IS THE AVAILABILITY RATE FOR SUBSTITUTES (%) ?
?65
HOW MANY DAYS ARE TO BE SIMULATED?  (MUST BE LESS THAN 190)
?180
THE RESULTS OF YOUR INPUT WILL PROBABLY TAKE MORE THAN 5 MINUTES
TO BE PROCESSED.  ARE YOU WILLING TO WAIT OR WOULD YOU RATHER
USE SMALLER NUMBERS FOR INPUT?  (WILLING TO WAIT=1, NEW INPUT=2)
```

Figure 5-27.
Second Run of SUBST Using 20 Substitutes. Continued

```
?1
PLEASE WAIT. SIMULATION IS ON THE 10TH DAY
STILL CALCULATING. SIMULATION IS ON THE 20TH DAY
PLEASE BE PATIENT. SIMULATION IS NOW ON THE 50TH DAY
SIMULATION FOR ALL DAYS COMPLETE. CALCULATION OF TABLE VALUES UNDER WAY
DONE WITH CALCULATING CUMULATIVE DAYS
DONE WITH CALCULATING % DAYS
DONE WITH CALCULATING CUMULATIVE % DAYS
```

# SUBSTS NEEDED LESS # SUBSTS AVAILABLE	# DAYS	% DAYS	CUMULATIVE # DAYS	CUMULATIVE % DAYS
1	2	1.1	2	1.1
0	2	1.1	4	2.2
-1	2	1.1	6	3.3
-2	3	1.6	9	5
-3	12	6.6	21	11.6
-4	11	6.1	32	17.7
-5	15	8.3	47	26.1
-6	23	12.7	70	38.8
-7	22	12.2	92	51.1
-8	17	9.4	109	60.5
-9	24	13.3	133	73.8
-10	17	9.4	150	83.3
-11	15	8.3	165	91.6
-12	8	4.4	173	96.1
-13	3	1.6	176	97.7
-14	1	.5	177	98.3
-15	2	1.1	179	99.4
-16	1	.5	180	100

Figure 5-27 Continued

REVIEW

Exercise

1. Make some runs of your own for the problem just discussed. Try 30 substitutes at least twice. Do the results meet the criterion set by the superintendent (not more than three days with less substitutes available than needed)? Investigate the possibility of fewer than 20 substitutes meeting the criterion. Can you find a lower limit for the number of substitutes needed to meet the criterion?

OVERVIEW OF SIMULATION

ADVANTAGES AND DISADVANTAGES OF SIMULATION

Now that you have a feel for what is involved in actual simulations, the advantages and disadvantages of the techniques should be summarized.

Simulation has these distinct advantages for the decision maker:

1. Simulation can be used in description, explanation, and prediction of the object system. It can also be used to teach about the subject system.
2. Simulation provides a framework for analyzing real-life situations into their components: we must identify the inputs to the system, the outputs from the system, and the rules by which inputs are related to outputs.
3. Simulation enables the decision maker to try out a variety of solutions on a model of a real-world situation without the time, danger, or expense of working with the real-world problem.

Some limitations of simulation are:

1. The object system under study must be expressable in the form of a model. This may result in an oversimplified version of the real-life problem.
2. Building models of real-life problem situations is often a complex and time-consuming task, requiring much technical expertise.
3. The predictions and solutions from a simulation problem are valuable only to the extent that the model for the real-life situation is valid.
4. Simulation provides guidelines for making decisions, but does not actually make decisions.

A FINAL WORD ABOUT COMPUTER SIMULATIONS

Most of the computer programs used in this book[12] are simulations. Computer simulations are particularly powerful tools for the educational administrator for a variety of specialized reasons:

- Computer simulations provide all the advantages of simulations in general.
- Computer simulations provide fast results at relatively little cost. The computer can generate solutions in minutes that would take people hours or even years to compute.
- There are usually programs available through computer centers and computer manufacturers. Hence, the time and technical expertise required to develop first a model and then a program can be shared by many users. This is an important factor in keeping the cost of computer simulation reasonable and affordable.

Before an administrator can use a computer simulation constructively, however, he must know something about the model used in the program. It is the user's responsibility to decide first whether the model is adequate for simulating the situation in question or whether it is too simplistic or unnecessarily complicated. To a large extent, the decision will depend on how the results are to be used. A model which yields adequate estimates for an administrator's personal projection figures, for example, may not be precise or comprehensive enough to provide figures on which to base a school bond initiative. If one is considering computer simulations as aids to decision making, then, the first thing he should decide is whether or not the model is adequate.

[12] GCPATH, QUEUE, LPRG, BUSRTE, ENROLL, and SUBST.

BIBLIOGRAPHY AND SUGGESTED REFERENCES

Barton, Richard F. *A Primer on Simulation and Gaming*. Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Maki, Daniel P., and Maynard Thompson. *Mathematical Models and Applications*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Mihram, G. Arthur. *Simulation, Statistical Foundations, and Methodology*. New York: Academic Press, 1972.

Naylor, Thomas H., Joseph L. Balintfy, Donald S. Burdick, and Kong Chu. *Computer Simulation Techniques*. New York: John Wiley and Sons, 1966.

Nelson, Robert D. "ENRPRO—An Enrollment Projection Program for K-12 School Districts." Richfield, Minnesota: Richfield Public Schools, 1973.

Reitman, Julius. *Computer Simulation Applications*. New York: John Wiley & Sons, Inc., 1971.

Shamblin, James E., and G. T. Stevens, Jr. *Operations Research: A Fundamental Approach*. New York: McGraw-Hill, 1974. Suggested reading: Chapter 7, Monte Carlo Simulation, pp. 156–186.

Shannon, Robert E. *Systems Simulation: The Art and Science*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Siemens, Nicolai, C. H. Marting, and Frank Greenwood. *Operations Research*. New York: The Free Press, 1973. Suggested reading: Chapter 6, Simulation Studies, pp. 153–174.

Springer, Clifford H., Robert E. Herlihy, and Robert I. Beggs. *Advanced Methods and Models.* Mathematics for Management Series, Vol. 2. Homewood, Illinois: Richard D. Irwin, Inc., 1965.

Van Dusseldorp, Ralph A., D. E. Richardson, and W. J. Foley. *Educational Decision-Making through Operations Research.* Boston: Allyn and Bacon, 1971.

INDEX