AMERICAN MANAGEMENT ASSOCIATION

# ADVANCED SCIENTIFIC TECHNIQUES FOR MANAGEMENT


Presented by the
Computer Systems Division
The Ramo-Wooldridge Corporation
Los Angeles, California

# PREFACE

The seven brief essays presented in this booklet form a part of the material to be covered during an Orientation Seminar of the American Management Association, to be conducted on December 12 - 14, 1956 and January 2 - 9, 1957. This Seminar, the fourth in a series on Operations Research also covers in a much more detailed way, such advanced techniques of Operations Research as Statistical Inventory Control, Queuing Theory, and the Monte Carlo Method.

# CONTENTS

# ELECTRONIC DATA PROCESSORS

by

Neal J. Dean
Head, Data Processing
Computer Systems Division
The Ramo-Wooldridge Corporation

## Introduction

Probably we should establish at the outset exactly what electronic data processors are and how they fit into a business operation. Many of you have probably heard the term EDPS which is frequently used to stand for Electronic Data Processing Systems. However, in our title we have used the word "Processors" rather than "Processing Systems." This is to deliberately emphasize the fact that these devices are really only tools; they do not in themselves constitute a complete business system. In an engineering sense, the combination of equipment that goes together to constitute an electronic data processing installation could be properly called a system; however, from the businessman's frame of reference, all of this equipment really represents only new tools for the implementation of a business system.

A business "system" is usually thought of as the basic structure by which the necessary outputs are obtained from the inputs; the methods and techniques for obtaining these outputs are referred to as "procedures." Since the electronic equipment falls within the category of tools for performing the operations specified by the system structure, they really do not replace a business system but merely mechanize parts of a business system.

Thinking of electronic data processing equipment as tools for the accomplishment of business operations, a systems and procedure man might then legitimately classify them with other tools such as the typewriters, adding machines, bookkeeping machines, and even pencil and paper. Of course, the electronic tools are much more powerful and offer tremendous possibilities which these more conventional tools do not. In this sense, the electronic data processing equipment might seriously influence the system structure; of course, conversely the business system structure does influence the particular data processing equipment selected to mechanize the structure. What, then, do people mean when they refer to the electronic data processing systems? It seems to the author that what is meant are data processing systems in which major portions have been implemented with electronic data processing equipment.

Recognizing that electronic data processors are simply tools for mechanizing business systems, let us consider what advantages these tools specifically offer over more conventional tools, then let us proceed to an examination of the specific functions of a business system that electronic data processing equipment can accomplish efficiently. We will then consider some of the specific equipment available and attempt to categorize it into its functional classifications. Next, we would like to point out some of the particular shortcomings in electronic data processing equipment, what is being done about them, and to prognosticate somewhat about future trends in this field. Finally, we have included a brief discussion of how a company might conduct an investigation of electronic data processing possibilities in its own applications.

## Advantages of Electronic Data Processing Equipment

First, to determine the advantages offered businessmen by electronic data processing equipment, we might logically begin by examining some of the characteristics of this equipment from which operational advantages might derive:

### High Speed

Electronic data processing equipment is capable of performing operations much more rapidly than the manual, mechanical, or electromechanical equipment presently used for business systems. As far as computations or logical operations themselves, these speeds are very high indeed, being in the order of 1,000 to 10,000 operations per second. In the realm of input-output speeds (for inserting the information into or obtaining information from a data processing system,) these rates are lower than the processing speeds but still very much higher than any of the other conventional equipment. For example, magnetic tape units can operate at speeds of approximately 15,000 characters per second and high-speed printers can produce reports at the rate of 1,000 lines per minute.

### Automatic Sequencing of Operations

The tremendous speeds quoted above would be of little value if it were not for the fact that sequential operations on electronic equipment can be performed automatically with human intervention. This is in contradistinction to punched card electromechanical equipment, in which the procedure for sequencing major operations is for an operator to take a deck of punched cards from one machine to another. In this way, punched card equipment requires that the total job be broken down into a number of distinguishable pieces with the link between these pieces being the human operator transporting cards; whereas, with electronic equipment the entire operation, in most applications, can be performed in one pass of the data through the machine.

## Increased Accuracy

The accuracy of an electronic data processing system is normally much greater than that of a manual, mechanical, or electromechanical system. This increased accuracy results partly from the large number of internal checking operations and error detection which can be performed in modern electronic data processing equipment. Also, because of the high speed mentioned above, it is also possible to "program" specific control checks during the course of a processing operation, thus introducing external controls in addition to the internal checking of the computer itself. Most of the increased accuracy, however, results from the great reduction of human intervention. Since the inherent accuracy of electronic equipment is many orders of magnitude greater than that of "human equipment," increased accuracy results directly from the substitution of the electronic equipment for human operators.

## Flexibility

Although one may get into quite a heated discussion as to whether an electronic computer is more flexible than a human being, a system in which electronic data processing equipment is used to a large extent is considerably more flexible than one in which human beings are combined with mechanical equipment. We are assuming here that the basic electronic computer is a stored program device, in which the operations it performs can be changed by simply changing the coded instructions, written to perform these operations.

On the other hand, the particular arrangement of men and machines, where mechanical or electromechanical machines are used to implement a specific system, is quite inflexible. Many times, even the machines themselves are special purpose, such as are some of the more elaborate bookkeeping or posting machines. Of course, the punched card equipment is considerably more flexible than such manually operated machines, but even here the particular combination of machines selected is on the basis of particular applications and their utility on applications which are quite different is usually not very efficient.

Now there are many benefits to business operations arising from these characteristics of electronic equipment. Among the important ones are the following:

## Reduced Cost

This largely arises from the increased speed and the automaticity of operations.

## Improved Reporting

The advantages here are really two-fold. One is that the reports are more up-to-date because they can be produced more rapidly than they can be under the manual or semi-automatic electromechanical data processing systems, while the second aspect of this improved reporting arises from the increased accuracy. Thus, more up-to-date and more accurate reports can be produced on electronic data processing equipment.

## The Consolidation of Files

This is a result of the rapid speed of processing as well as the automaticity of operations, making it possible to perform a much greater proportion of the total job in one pass of the information through the processing equipment than is possible on conventional systems. This consolidation of files is a virtue both because of the space saving resulting from it as well as the increased accuracy which it makes possible. The latter follows directly from the number of files, thus making it unnecessary to be sure that a large number of files are in agreement for common information; for example, it is not necessary to be sure to post any changes to many different files.

## The Automatic Processing of Exceptions

To a much greater extent than is ever possible on punched card equipment, it is feasible with electronic devices to process almost all of the conceivable exceptions that might arise in data processing. This results from a combination of the speed, the automatic operations and the flexibility of the electronic equipment.

## Data Processing Functions

We would next like to discuss some of the data processing functions which can be efficiently mechanized with the electronic tools. These include the following:

### The Transcription of Information

This refers to the basic transcription of data. Business information may be produced in a number of ways as a result of events occurring in the business system, but it should be recorded as soon as possible in a form which is machine readable; that is, capable of being read directly into the data processing equipment.

### Transmission

Data processing equipment can be used for the transmission of information from one location to another. Radio, telephone, teletype or television may be used for transmitting coded information rapidly from one geographical location to another.

### Storage

Information can be stored in many ways in electronic equipment. Usually it is recorded in the form of coded data representing, for example, a historical or other reference file, a current file of transactions, information required for a report, or the instructions required to execute the program.

### Processing

Processing may be divided into two major categories:

#### Computations.

These include the arithmetical operations of addition, subtraction, multiplication and division, and their combinations for the solutions of more complicated scientific problems.

Logical Operations.

These include transfers, comparisons, selection of the lower of two quantities, etc. A sorting or sequencing operation utilizes logical operations for arranging units of information in accordance with a specified sorting key; the logical operations required include comparisons, transfers, and "branching" operations (in which alternative routines are selected on the basis of comparisons). These branching operations are essential to the very important "decision making" ability of the data processor. Similar operations are required for classifying data for statistical distributions.

## Equipment

Let us now consider some of the equipment used as tools for accomplishing these functions.

### Transcription Equipment

This is the basic equipment by which it is possible to record information in a machine-readable form. Various media which can be utilized in this operation, and some of the associated devices and their characteristics are listed below:

Card Punches.

By manually operating key punches it is possible to record information in punched cards. The IBM 80 column cards and the Remington Rand 90 column cards have been used for many years. In addition, a newer development--magnetic cards--makes it possible to record information by magnetized spots rather than punched holes in cards. This increases the information capacity of a card very significantly.

Punched Paper Tape.

By manual key operations, it is possible to perforate coded information in paper tape. This can frequently be done as a by-product of preparing documents on a keyboard device. The Flexowriter and various paper tape attachments to adding machines and bookkeeping machines are typical of the devices available in this field.

Magnetic Tape Preparation Unit.

There is now available a device for recording information directly on magnetic tape. This device is referred to as a Unityper and is produced by Remington Rand to operate with their UNIVAC equipment.

The above are the major categories of devices for preparing input media. In addition, there is a need for verifiers which verify the accuracy of the coded information in various media. In this case, a second operator inputs the same information on a similar device which compares the original coded information with the second operator's key strokes detecting any deviation between the two sets of information.

## Communication Equipment

Transceivers.

This is an IBM device which makes it possible to transmit information coded in the 80 column cards from one location to another by means of either telegraph or telephone lines or even by radio channels.

Punched Paper Tape Transmitters.

There is a large variety of equipment which will transmit punched paper tape information over telegraph or telephone lines. These devices are made by several companies including Teletype Corporation, and can utilize standard Western Union and AT and T transmission facilities.

## Information Conversion Equipment

It is frequently necessary to change information from one form or media to another in a complete data processing system. Such conversions should, of course, be minimized, but in many cases they cannot be avoided. Some of the equipment available for conversion include the following:

Card-to-Magnetic-Tape Converters.

These devices convert the information from punched cards to magnetic tape, in order that the speed of input to the computer can be increased beyond that possible if punched cards were to be used directly.

Magnetic Tape-to-Card.

This equipment performs a similar conversion on the output of a computer where magnetic tape has been produced directly and it is desirable to have this information in punched card form.

Punched Paper Tape to Magnetic Tape or Magnetic Tape to Punched Paper Tape.

There are devices available which make conversion between these two tape media possible.

## Storage Equipment

Storage of information can be accomplished on a number of devices; these include magnetic tapes, magnetic drums, magnetic discs, punched cards and several new optical devices, whereby it is possible to store information on film or by very small photographs on "minicards."

## Processing Equipment

The equipment which is used for actually performing the computations and logical operations on the information in the various media mentioned above include the following basic types:

Punched Card.

There are a large number of devices available for this purpose in the punched card field including collators, sorters, tabulators and computers.

Electronic Computers.

In the electronic data processing equipment, the computers themselves are capable of performing all of the collation, sorting and tabulating operations performed by the individual units in the punched card field.

## Output Equipment

Although the direct output of the computer may be magnetic tape, punched paper tape or punched cards, the ultimate output of a system should be information which is directly intelligible to human beings and, therefore, in printed form. There are a large variety of printers available, including the following main categories:

### Line Printers.

There are high-speed line printers available which will operate at a thousand lines per minute or higher. The more conventional tabulators operate at one hundred or one hundred and fifty lines per minute.

### Typewriters.

Typewriters prepare printed copy at the rate of a character at a time. Speed of printing here is in the order of ten characters per second.

### Plotters.

Plotters are available for presenting output information in graphical form rather than in tabulated lists.

## Electronic Computers

The electronic computers themselves may be classified according to several different characteristics; two of the most important probably are their purpose and their size.

There are two general classifications as to utility, general purpose and special purpose. The general purpose type is, of course, much more flexible. Computers of the general purpose type are of the "stored program" variety, whereby the individual instructions which succeed in executing a processing operation are coded and recorded in the storage unit of the computer. Special purpose devices sometimes have some stored program features but are usually of the fixed-program variety, whereby the instructions are actually wired into the machine.

A general purpose computer is designed to handle a large variety of applications, whereas a special purpose computer is designed to handle a single or very restricted type of application. The use of a special purpose computer thus implies that the application for which it is intended is of sufficient size to justify the device for this application alone.

Probably the most widely known special purpose computer is the Magnetronic Reservations System for the airlines. Another is ERMA, a system designed specifically for deposit accounting for commercial banks, presently being put in production by General Electric Company.

Now from another viewpoint, computers may be divided according to their size into large, medium or small scale. Arbitrarily, these may be categorized by their price. Computers with a purchase price (or equivalent rental) of about $1,000,000 or more can be considered large scale; those costing in the order of $100,000 to $700,000 considered medium scale; and those that sell for less than $100,000 as small scale.

The computers themselves include devices which make it possible to read information in (input devices); units for storing the information (internal storage or memory unit); for performing the necessary computation and manipulation of data (arithmetic and logical unit); for producing the various outputs required (output devices); and for exercising the necessary functional control of the various units (control unit).

We will not discuss the arithmetic and logical unit or the control unit because to do so would involve quite a bit of technical detail which is not essential for our present purpose. Also, the output devices have been covered above. However, both the input devices and the storage units have some features which are important to consider with the electronic computer itself, and some of these characteristics are listed below.

## Input Devices

Devices for inserting information directly into the computer include the following:

1. Paper tape readers which operate in the order of 60 to 600 characters per second. Some devices operating as high as 1,000 characters per second have been demonstrated on a laboratory basis.

2. Punched cards can be read at speeds of 250 cards per minute. (One of the large-scale systems to be delivered in the near future reads cards at the rate of 900 cards per minute.)

3. Magnetic tapes in which reading speeds are from 1,000 to 60,000 characters per second are available.

## Storage Devices

Two basic characteristics of storage devices which are of importance in the data processing system are capacity and access time. The capacity refers to the total amount of information which can be stored in the device and the access time refers to the period required between the initiation of a request for information and the delivery of that information.

There are two access times which are of significance. One is the random access time and the other, the sequential access time. The former refers to the period of time required to obtain the information when the requests are in random order. The latter refers to the period of time required to obtain information on two successive file entries from the storage device to contiguous file records.

Let us briefly consider the orders of magnitude of both the capacity and the access time for typical storage devices:

### Magnetic Cores.

The access time here is the same regardless of whether random or sequential access is required, and is in the order of ten to twenty microseconds (a microsecond is one millionth of a second) per character. Electrostatic storage tubes (cathode ray tubes) and vacuum tube storage units have access times in the same order of magnitude as the magnetic cores. However, they are both considerably more bulky than the core devices. Capacities of magnetic core storage units are in the order of 200 to 20,000 characters.

### Magnetic Drums.

Magnetic drums might be referred to as a quasi-random access in the sense that any track can be directly addressed, but within the track we must wait until the drum rotates to the proper record of information. However, these are frequently referred to as random access devices.

The random access time for magnetic drums depends upon the speed of the drum and the speed of the addressing circuitry, but is in the order of 1/60th of a second. The serial or sequential access time (from one piece of information to an adjacent piece of information on the same track, i.e., the next record) again is in the order of a few microseconds per character. Typical capacities are 1,000 to 180,000 characters. Some novel drums have been developed with capacities greater than one million characters.

## Magnetic Disks.

IBM has recently announced a quasi-random access device referred to as the Random Access Memory, consisting of a "juke box" type of storage utilizing magnetic disks. An arm moves up along a vertical stack of disks and then moves into the proper track on the disk. The device appears somewhat similar to a Wurlitzer Record Player. The average random access time is in the order of 1/2 second; the serial access time is again in the order of microseconds. The capacity of this memory device is five million characters.

## Magnetic Tapes and Punched Paper Tapes.

Both of these devices are of the sequential access type. The serial access time depends upon the particular unit. For magnetic tapes, typical reading speeds are from 1,000 to 60,000 characters per second; thus, the serial access time would be in the order of 16 to 100 microseconds per character. Punched paper tape operates at a much lower density and speed. Typical readings speeds are from 60 to 600 characters per second, yielding serial access times in the order of 1.6 to 16 milliseconds (a millisecond is a thousandth part of a second) per character. The capacity of a 1,000 foot reel of paper tape would be 120,000 characters ( 10 per inch).

The random access time depends upon the length of tape used. It really is of very little significance for punched paper tape, since they are rarely, if ever, used as a random access memory. For magnetic tapes the random access time corresponds to the length of time required to traverse through one-third of the tape (not one-half as might be expected) from any one random position to any other random position on the tape if the tape can be moved in either direction and programmed to go in the proper direction for the particular record desired. For a fifteen hundred foot reel of tape operating at 60" per second, this would correspond to about 100 seconds for the average random access time. Capacities range from one to five million characters per reel of magnetic tape.

## Installations

Now that we have discussed what electronic data processors are, what functions of a data processing system they can conveniently mechanize, and described briefly some of the specific pieces of equipment that can be used, and how they are categorized, it might be well to quickly summarize the specific accomplishments in the business data processing field to date. There are over 600 installations of large, or medium scale data processing systems working on business applications in the country today. About 75 of these are of the large scale type (IBM 702 or 705, UNIVAC I or RCA BIZMAC). The range of applications have extended from the Government to practically every line of industry and business, including both industrial production and such service organizations as insurance companies and banks.

I don't think there is sufficient data at the present time on actual operations with these business data processing installations to accurately assess their success. One thing that can be stated safely is that the debugging period has exceeded expectations in most cases. As might be expected, those companies which did detailed planning and programming, and even attempted to debug their programs on the computer manufacturer's service bureau facilities, did much better in getting their computers into actual productive work than did those companies which did a less complete job of preparation.

## Recent Developments and Future Trends in Equipment

It might be interesting at this time to consider some of the shortcomings of data processing equipment as practical business tools, and what one might expect in the way of future attempts to overcome these deficiencies. First of all, one of the more awkward operations on data processing equipment (in comparison to the same operation on punched card equipment) is sorting. In many cases it proves more economical to perform the sorting on the punched cards before converting them to magnetic tape rather than to do the sorting on the computer itself. Especially is this true of those electronic data processing systems which include an auxiliary off-line data processor specifically for sorting, sequencing, extraction, merging, etc. In these cases sorting can frequently be accomplished much more expeditiously and, even more importantly, without tying up the main computer itself. These auxiliary devices offer a great deal of promise for improving the performance of electronic data processing equipment in this area.

Another deficiency -- in the magnetic tape equipment at least -- is the lack of sufficiently rapid random access for either answering inquiries, where there is a large volume of such interrogations, or for processing in an on-line fashion. An on-line system is one in which the event is immediately reflected in the basic file of information by inputting the data directly and updating the file at the time the event occurs. Such on-line systems, where the inputs are in random fashion (as they are in most applications), are difficult to accommodate with serial access memory devices. However, there have been a number of recent developments in "random access" memories. One of the more important ones was mentioned above in the storage section--the IBM "juke box" memory device. There have also been some laboratory investigations of the possibility of accumulating a number of random access entries for a short period (a tolerable period for the specific system), sorting the transactions and then processing them in sequence by scanning through the entire file. The equipment for this operation has not been fully developed, however, and specific time requirements do not appear to have been sufficiently well surveyed to determine specifications of such equipment accurately.

Another area receiving a great deal of attention is the output area where higher and higher speed printers are being developed. Some of the more novel devices make it possible to produce a whole page at a time by displaying the information on a cathode ray tube (similar to a television tube) photographing it on microfilm or 35 mm film, processing the film, and finally converting it to "hard copy" on an off-line basis.

However, it seems to us that some of this frantic search for higher and higher speed outputs is somewhat misguided, and that in many cases, the real solution to this problem is to reduce the amount of output by making the computer perform more of the basic decisions which lower level management would make after scanning the output reports. In other words, a successful implementation of management by exception should considerably reduce the output requirements. Of course, this will not be true of all applications, especially where such documents as bills, statements, etc., are produced for transmission to the customers.

## Company Approach to EDPM Investigations

Perhaps no discussion of this electronic data processing field for business applications would be complete without a few remarks on how a company might reasonably approach an investigation program of the electronic data processing possibilities for their own applications. One good way is to conduct a feasibility study in those areas which seem to be the most attractive.

It appears that the best way to approach this is to appoint a high-level committee, representing a complete cross-section of the entire business operation, for the purpose of organizing and reviewing the progress of such investigations. One of the functions of such a committee would be to appoint a study team. This study team might consist of present employees of the company, or the study could be conducted by outside consultants.

Without considering in detail the advantages and disadvantages of both approaches, some of the more pertinent considerations are the following. To conduct such a study successfully requires a very high level of talent. In addition to a detailed knowledge of the present applications, a very imaginative type of mind is required -- one which is capable of breaking away from present procedure restrictions to consider possible innovations. Also required is a knowledge and understanding of the equipment available in this field, and the types of systems within which this equipment operates efficiently. The first requirement for a study team can usually be met by employees of the company; to find this detailed knowledge of the present system in combination with the other requirements, however, is a rarity. It seems to us, therefore, that in many cases a team effort is required.

Even if such characteristics can be found within the employees of a given company, such individuals are usually key operating executives or supervisors within the company and cannot be spared on a full-time basis for such an investigation. (It should be emphasized here that it appears absolutely essential that these individuals be available on a full-time basis. Proper justice cannot be given to either job is such an arrangement is not made.) In this case it may prove an important advantage to the company to obtain such help from outside consultants. The consultant will require assistance from the client's operating people, especially in the area of problem definition -- supplying the details about present operating procedures. However, the over-all direction of the investigation and the technical systems and equipment considerations can be adequately supplied by such outside experts.

Concurrently with such detailed investigations, it appears that a company should embark upon an education program, to acquaint its executives and operating personnel with the basic characteristics and potentialities of electronic data processing equipment. This is important, not only for the purpose of properly directing the investigations, but to expedite the ultimate conversion if the investigations are successful.

Detailed feasibility studies should be conducted in those areas which appear to be most attractive. To determine those which are most attractive, sometimes really requires a detailed study. However, such detailed studies in all areas usually would require so much time that an opportunity for very large savings can frequently be missed. Therefore, perhaps the early investigations should be of a survey type, to quickly determine the areas which seem to offer the greatest opportunities. Some characteristics can be helpful in recognizing such promising areas, such as a large volume of repetitive operations, particularly where the operations are well defined.

Once the areas have been determined for the detailed investigations, the feasibility study proper should be commenced. This should begin with a detailed investigation of the present procedures, quantified as to volumes, frequency, time schedules, documents, input and output requirements and communications. The team should then consider how such a system (or deviations therefrom which will still accomplish the same result) can be mechanized with electronic data processing equipment. A general data processing system should be outlined and flow charted, indicating the general framework within which this application will be accomplished on the equipment.

Simultaneously with these feasibility investigations should be conducted a survey of available equipment, in order that the general system outline can be implemented with the existing equipment. A specific comparison of the abilities of the individual equipment to meet the requirements of the general data processing system outlined should then be investigated. This comparison should include both economic and operational considerations. As a result of these comparisons, it should be possible to arrive at a particular combination of commercially available equipment which best "mechanizes the data processing system" determined as described above.

After the results of this investigation have been properly reviewed by the committee, and their accuracy verified particularly as to potential cost savings, a decision should be reached as to whether this is a good application for electronic data processing equipment of the type outlined or not. If so, presumably the equipment would be ordered, and the detailed planning for the installation and conversion commenced. This phase would include the following: over-all systems layout; programming and coding the application; the design of the required forms; the determination of the detailed input requirements and a design of the input source documents; specifications of the reports required and the report formats; detailed time schedules for the particular portions of the application to be performed; organization layout and detailed conversion plans from the present system to the electronic system.

Finally, with the arrival of the computer itself, the installation phase begins. This will include performance and acceptance tests, the implementation of a conversion program, the debugging of the individual instructions -- the programs written for the specific applications and the over-all systems tests and "shake-down" on the actual application. It may also include some transition procedures, and it may be necessary in some applications to conduct dual operations, both on the old system and the new, for a short period of time. It should be emphasized that it appears reasonable to convert from the old system to the new system gradually, taking a certain proportion of the work in certain periods of time, so that a general disruption of the entire application will not result.

In many cases, it will develop that no one application can justify a computer -- even a medium scale computer. In other cases, if a single application is considered, a medium scale computer might prove the best; whereas, that application in combination with others, would perhaps result in a more efficient system if placed on a single large-scale computer. Thus, there is the problem of combining applications to determine the true economic feasibility.

However, this is frequently necessary anyway, because of the lack of independence between applications or operations within a company's over-all business. This might be referred to as the integration consideration. For example, in a production control or inventory control application, such input source documents as receipts, orders, and material requirements are actually the results of other operations such as receiving, purchasing, sales, and order processing, and are intimately connected with accounting operations such as accounts payable and receivable, invoicing, and the preparation of shipping documents. We see that all of these operations are closely interrelated. This implies, and it is indeed true, that a truly successful approach to the solution of these problems must include the all-important integration consideration; that is, the outputs of one operation should be compatible with the input for the next operation. In addition, the over-all time schedule should be in agreement; this frequently requires compromises between the various operations. It is, therefore, a consideration which is, in most cases, absolutely essential for the successful implementation of a data processing system.

## Summary

To summarize, in this brief essay we have attempted to outline briefly the following:

1. what electronic data processing equipment really is -- that it is merely another tool for the implementation of a business operating system;

2. indicate specifically what functions of a business operating system one might expect electronic data processing to successfully mechanize;

3. consider the various equipments available for the mechanization of these functions;

4. indicate some of the more recent developments in the field and what might reasonably be expected in the near future;

5. finally, we attempted to outline very briefly, a possible approach to the investigation of electronic data processing applications to a particular company's problems.

# AUTOMATION

by

E. M. Grabbe
Senior Staff Consultant on Automation
Computer Systems Division
The Ramo-Wooldridge Corporation

## Introduction

Automation may be defined as the use of machines that are self-powered, self-guiding, and self-correcting to perform physical and mental jobs, many of which were formerly done by man. These devices extend and replace human effort and senses. The object is to increase our output per worker, reduce production costs, and improve the quality of products.

There is no automation industry, as such. Manufacturers of automation equipment, as well as its users, cut broadly across all fields of business and industry. A convenient classification of the elements of automation is (1) advanced mechanization, (2) feedback control, (3) electronic computers, and (4) systems design approach. The typical automation system will contain all of these elements in various proportions.

Most of the advances that have led to our high productivity of today have been on the basis of advanced mechanization and feedback control. Extensive employment of the third and fourth items, computers and systems design, is relatively new. While people speak of today as the age of electronics, it is important to note that most of our accomplishments in automation to date have come about with relatively little use of electronics, as shown in Table I. Electronic computers, controls and measuring devices will make many inroads into presents operations. Systems design will also play a more important role as automation applications become larger and more complex. Previous emphasis has always been on individual components.

Table I. Automation Examples

| Application | Techniques of Today | Techniques of Tomorrow |
|---|---|---|
| Chemical and Petroleum | Pneumatic Controls | Electronic Controls and Electronic Computers |
| Automotive | Mechanical<br>Electrical<br>Hydraulic<br>Pneumatic | Electronic Gauging and Control will be added. |

Let us briefly summarize the status of the four elements of automation, and then evaluate recent progress in automation and its impact on management.

## Advanced Mechanization

Increased mechanization means larger, faster, and more complex machines with automatic handling of materials between machines. On a recent television show Groucho Marx referred to these machines as, "the nuts and bolts with a high IQ". Typical examples are (1) motor block machining in the automotive industry, (2) television picture tube production, (3) aircraft skin-millers, and (4) automatic assembly of electronic components for radio and TV.

Experience to date indicates that these large machines are expensive but if properly employed will pay off by cutting costs. Applications generally involve a high volume of a fixed product.

There is no general formula for applying automation of this type but the greatest savings are obtained when radically new approaches to production processes are employed, using new machines, new materials, and radical designs. One great benefit of trying new things is that frequently unexpected benefits occur from new methods. For example: radio tubes designed for automatic production have better physical and electrical characteristics.

## Feedback or Self-Regulating Control

Modern technology has provided many instruments that make it possible to measure continuously the condition of the process. Then if deviations occur this information is used to control the process to bring it back to some desired condition. That is what is referred to as self-regulation or feedback control. Many modern processes that require close quality control have been speeded up so that today man can no longer make measurements and take corrective actions quickly enough and automatic controls are essential for operation. Examples are: (1) the modern complex chemical and petroleum processing plants, (2) steel strip mills, (3) continuous processes in the paper industry using nuclear gauging instruments, and (4) automatic production gauging of ball-bearings, gears, and other small precision parts.

Feedback control is used where it is necessary to control closely the specifications of a product. Without feedback control high speed operation can produce useless scrap at the same high speed instead of a uniform valuable product.

Much of the present-day feedback control instrumentation was developed by practical people without modern theory. The theory of servomechanisms developed for gunsights, radar and other military applications during the last ten to twelve years can lead to many improvements both in design of new plants and operation of existing plants.

We can look for some spectacular advances in use of new measuring instruments and new designs of equipment and controls on the basis of theory and system studies. These developments will proceed slowly for equipment cost is high, experimentation is expensive, and good control engineers are scarce.

## Electronic Computers

The electronic computer is simply a device, such as a desk calculator, for carrying out mathematical calculations; but the electronic computer works at a much higher speed, perhaps fifty to a hundred thousand times faster. It can be expected that the computer will become a very important factor in our economy taking over much of the information processing activity handled by human beings and completely revolutionizing the activities of the workers of the industrialized part of the world. The use of computers to take over clerical work is fairly obvious. However, we should note that a large part of the activity of all industrial workers is concerned with processing information and providing the intelligence to operate machines and tools associated with translating product specifications into a useful product.

The power of electronic computers lies in their high speeds of operation, their large storage capacities, and their internally stored programs which make possible for them to carry out long sequences of operations without the intervention of humans. Computers are extremely versatile and may be used for scientific and engineering computation, business data processing, for control of processes, and for logical decision processes.

Logical decision processes are concerned with the ability of the electronic computer to make decisions. Using this ability, computers can be applied to language translation, library referencing, medical diagnosis, and many other applications. Recently on NBC's stunt show "People are Funny", a Univac Computer, attempting to remove the haphazard factor in marriage, sifted through information on 4,000 couples to select an ideal match. Latest reports are that the engagement has been announced. Computer applications of this type eventually may far exceed their applications as arithmetic computers.

Many benefits have come to business from the use of electronic data processing systems. These benefits include: (1) increased speed, (2) higher accuracy, (3) new information for management, (4) the use of multiple inputs and outputs in business problems, (5) consolidation of files, (6) automatic processing of exceptions, and (7) integration of systems. Cost benefits accrue from reduced clerical staff, savings in floor space, reduced inventory, and other factors depending on the application.

It is important for management to understand what computers can do for business and what their benefits are. This is a new and powerful tool, and management cannot afford to relegate it to the role of the accounting machine.

## Systems Engineering

To gain the greatest benefits of automation in business, industrial and military application, a thorough analysis of the system must precede application. This means placing emphasis on the complex pattern of men, machines, materials, methods, and money -- the "five m's" that constitute a control system. Previous emphasis has always been on the individual components of such a system.

The systems approach focuses attention upon the function and purposes of a system. It brings to light alternate ways of doing a job, simplified design or new design of products, and many new products. For example, the acceptance of a product may be determined by the method of packaging, which reflects back into production requirements. In our defense programs, the workability of complex military systems depends on good systems engineering.

Systems engineering requires a broad background in mathematics and science. An outline of the steps involved in applying systems engineering are:

1. Formulation of the problem.
2. Analysis of functions.
3. Block diagram.
4. Detailed system design.
5. Detailed equipment design.
6. Equipment and system test.
7. Final design.

A good statement of system design that emphasizes the importance of theory is that today one must analyse a system before building, while building, and during test and evaluation.

## Automation and Management

To cope with the new problems of automation, management has recognized the necessity for more technical understanding. This does not mean knowing the details of operation, but rather understanding what the benefits may be. To illustrate how companies in diverse fields are carrying out educational programs, let us take several examples.

The General Electric Company has a nine-week management school at which the managers are brought up to date on new management techniques and new technology. Two-hundred fifty management personnel attend these nine-week sessions each year. The plant managers of tomorrow must be more technical because, with automation, production engineering is elevated to a new position. At General Electric, for example, there are now more graduate engineers involved in production engineering than in design engineering.

At the other extreme in the application of automation to finance problems, the Bankers Trust Company of New York has employed The Ramo-Wooldridge Corporation as a consultant in electronic data processing. As part of this program a course on electronic data processing was given to forty persons from operating management. This course consisted of twelve two-hour lectures over a six weeks' period. A shorter, more concentrated course was given to one-hundred twenty persons from the senior executive level, and plans are being made to extend this to several hundred executives.

With greater emphasis on production engineering, this means that more talent is being applied at the production level. The cost of personnel will be higher and the equipment they design will also be very expensive. Since production is planned for a product on the basis of anticipated demand, it becomes the responsibility of the market analysis and sales forecasting activities of a company to determine what this demand will be. Tooling up for low volume production will be quite different from high volume production. It can be just as costly to a company to underestimate demand as it can be to overestimate demand. New techniques are needed for better sales forecasting. All this is part of the systems picture. It means a close tie in between engineering, production, and sales. The techniques of operations research must be employed to their fullest extent.

The import of electronic digital computers is being felt most of all in en-engineering design. Design by model building and manual computation can frequently be by-passed with a modern electronic computer. One large company which manufactured transformers required an average of fourteen weeks to design a transformer for production. By standardization of parts and use of a computer for calculation, this design period is now cut to three days with a total computer time of an hour or so. With standardized parts, the transformers can also be produced more quickly so that the period from order to delivery can be cut possibly by 50 per cent to 75 per cent. Competition for business will be keener than ever since such drastic improvements in delivery schedule can attract additional business. Companies, instead of being equally inefficient with long delivery schedules and frequent delays, will now compete on an entirely new basis set by automation.

## Conclusions

1. Management can look for a period of accelerated technical developments and broader application of automation techniques.

2. Management must understand and accept new technology applied to production and business.

3. In applying automation the systems concept must be understood and applied. Operations research and long-range planning must be emphasized.

4. Marketing, sales and advertising become closely related to automation in that the automation tooling for a given product will depend on the volume anticipated.

5. The electronic computer will link production with the office, for the same tools are employed in engineering design, production scheduling, product design, and market forecasting. The importance of the computer in automation makes it especially important that management keep abreast of new developments in this field.

# SERVOMECHANISMS AND FEEDBACK AMPLIFIERS

by

E. M. Grabbe
Senior Staff Consultant on Automation
Computer Systems Division
The Ramo-Wooldridge Corporation

## Introduction

Modern technology has provided many instruments that make it possible to measure continuously the conditions of a process. Then if deviations from some desired condition occur, this information is used to control the process to bring it back to this desired condition. This technique has come to be known as self-regulation or feedback control. The devices that produce the action are called servomechanisms or feedback amplifiers. Historically, feedback control was first applied to electronic tube circuits to improve their stability. During World War II the theory and practice of servomechanisms were developed in military systems for directing gunfire. The controlled variable was a mechanical position. In both mechanical and electronic devices of this type, amplification is an essential part of the equipment since usually some useful task is accomplished.

An example of a feedback control amplifier is the automatic volume control now used on all radio sets. (You may recall the early days of radio when a weak station would fade in and out during the course of a program.) An automatic volume control constantly examines the signal. If it is too weak the amplification (gain) is increased, and if it is too strong the amplification is cut back. Hence, a desired volume level is attained.

An example of a positioning servomechanism is the ride control used on one modern automobile. As the passenger load in the car is increased a motor drive automatically adjusts the tension of the spring suspension so that the correct riding comfort is obtained.

The reason we cannot sharply distinguish between servomechanisms and feedback control amplifiers is that in many cases the techniques overlap. For example, Figure 1 shows an automatic curve tracer. A curve mask is placed on the face of a cathode ray tube and an electron beam is made to follow the curve by photoelectric measurements. If the beam is too high, the photoelectric tube received a large amount of light and will signal to reduce the deflection voltage. If the beam falls below the mask, the photoelectric intensity drops very low and the beam is again repositioned to ride the top of the curve. The beam spot is positioned electronically.

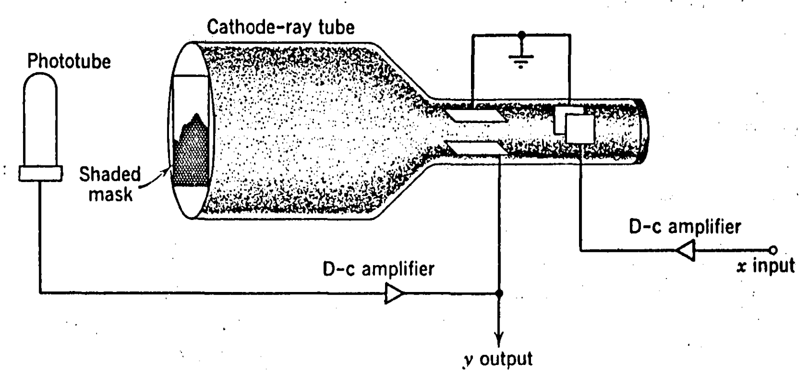



Figure 1. Electronic Curve Tracer

The output is a voltage representing the curve portrayed by the mask.

## Feedback in Business and Industry

Historically, business processes involve long time lags and frequently offer difficulties in making quantitative measurements--adverse conditions for good feedback control. With the advent of electronic data processors and the possibility of "in-line type operations", the ideas of feedback control must be re-examined by business in developing new systems.

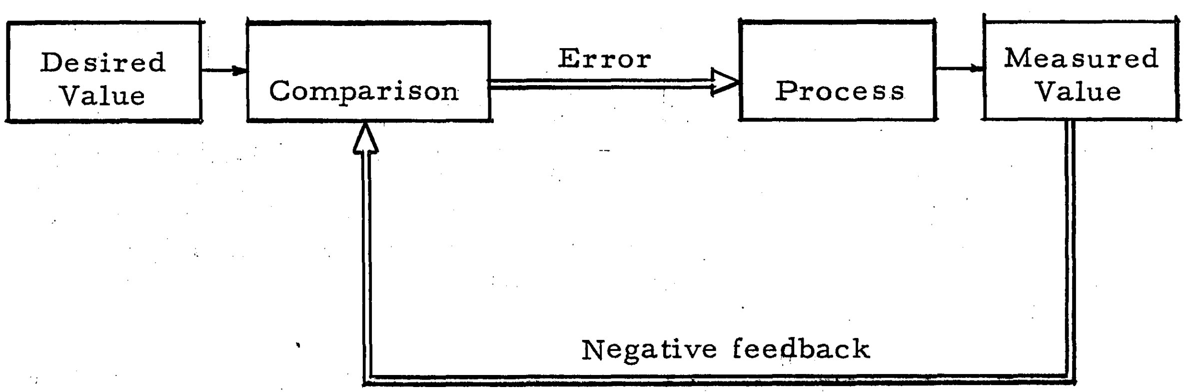
In discussing the application of feedback control to management, a one-to-one parallel may be drawn with a process control system. The control steps are shown in Table 1.

Table 1. Steps in a Feedback System

| | Example: Business Inventory Control | Example: Industry Annealing Furnace |
| --- | --- | --- |
| Step 1. Measure the condition of the process. | Inventory count | Temperature |
| Step 2. Compare the measured value (feedback) with a desired condition and determine the difference (+ or -). | Count versus forecast | Measured versus desired |
| Step 3. Apply corrective action in time to attain the desired condition. | Reorder if inventory low. Change policy if inventory too high. | Increase heat if temperature low; decrease if too high. |

The above examples are simple but they illustrate that feedback control is an essential part of all purposeful human activity. The sequence may be illustrated in Figure 2 which shows how the loop is closed.

Figure 2. Feedback Control



Negative feedback means that the measured value is subtracted from the desired value to obtain a correction signal (error). The correction will always oppose an undesired process trend as indicated in Step 3, Table 1.

## Closing the Loop

It is difficult to imagine many control systems that have no feedback, i.e., are "open loop". This would mean that something would be set into operation with no one caring anything about the results.

If the loop is closed automatically, we have a servomechanism or feedback amplifier. In many cases, the human closes the loop rather than a machine. For example, piloted aircraft versus auto-pilot control, automatic washer versus manually controlled self-powered washer. In business, humans usually close the loop. There is usually a time lag between the measurement and the control action. In many cases, this leads to instability in control.

## Stability

To be useful a control system must be stable. When a feedback is used there is always the danger of instability or poor quality of control. This can come about from having positive feedback instead of negative feedback, or it may be due to time lags in the system.

## Positive Feedback

A very simple example of instability is the dual-control electric blanket in which the controls for husband and wife have been switched. As soon as the wife turns up her blanket to become warmer, her husband's blanket becomes warmer. He immediately turns down his control and his wife becomes cooler. The result is what is called a runaway action because the feedback is positive rather than negative. In a chemical or petroleum plant this might lead to a fire or an explosion. In a manufacturing plant it can cause a lot of scrap.

Another simple example of an unstable system is a microphone which picks up too much of the amplified output and goes into a resonant howl. In business, positive feedback would mean going against common sense --for example, increasing production when sales lag.

## Time Lags

Stability of a system is frequently associated with the time lags involved in applying corrective action. If the corrected action is $180^{o}$ out of phase, as in the case of the electric blanket, it will act as a positive feedback and produce a great instability.

Automatic control used in industrial systems have time lags of the order of fractions of seconds, up to fifteen minutes. In systems with longer time constants involving days, weeks, or even months it is probably impractical to use automatic control. In business, long time lags are inherent, hence, it is likely that for some time to come in business applications, feedback will appear through the human operator.

Any business or industrial operation which has a purpose must employ a form of feedback. If a new product is marketed, customer acceptance is important, reliability of the product, its durability, the trends in marketing, all of these are elements based on what has happened which will modify the future activities of a business and, hence, correspond to closing the loop in a feedback system. With the advent of automation in both the factory and office, the speeding up of production and data processing will result in shorter time lags in closing the loop. We know from experience that if data is too old, it is possible worse than no information at all for it may be so much out of date and incorrect that it leads to improper actions.

## Cybernetics

Today we hear much talk of feedback control in discussions of cybernetics and the human nervous system. This is no accident, for modern machines and men have feedback control in common. Without this, it would be impossible for machines to take over many of the tasks formerly done by man. Practically all human physical movements employ feedback loops which signal our muscles to indicate that some desired action is being successfully accomplished. Certain diseases that impair these feedback loops of the nervous system in man give rise to oscillatory motions and the inability to carry out even simple manual operations.

## Forecasting

The trend of the future will be to predict future behavior. Prediction or forecasting also involves feedback. Again, humans and animals exhibit this ability to predict in a feedback control system. Prediction is involved in driving an automobile, and the hound chasing a fox will not run toward the present position of the fox but will attempt to head him off by running toward some point where the fox will be at a later time. This is similar to the modern interceptor fire control system which must direct a gun at some future target position in order to score a hit.

In business, the emphasis must also be on forecasting for the future. There will be pressure to obtain even more up-to-date information to keep ahead of competitors and to utilize available information as effectively as possible. Hence, we can see that feedback is closely related to automation and information theory.

## Implementation

It is very well to talk about using feedback control in business, but general discussions of the principles and emphasis of their existence do not show management how to use these principles. In engineering, accurate measurements are possible and theoretical analysis of the performance of a system can be made that relate the input to the output of a system. The mathematical expression relating input to output is called a transfer function. Knowing the transfer function of a system, it is possible to predict performance and stability of control. Finally, criteria for optimizing control can be specified.

One can measure the way a wind is blowing by holding up a wet finger. But engineering instruments are available also for accurately measuring the direction and magnitude of the wind. In business, the problems of measurement and analysis are much more difficult. Trends can be frequently assessed but accurate forecasting of the details is difficult. Electronic computers will find broad application to this problem.

Many industrial control problems involve exact mathematical relations between several variables. The relation of business performance to management action, on the other hand, has more influencing factors, many of which are difficult to define and measure. These factors range from government tax policies, labor relations, and the international situation to seasonal trends, population shifts, and habits of the individual. The advent of the electronic computer makes it possible for both business and industry to solve complex problems involving many variables. Examples are (1) the SAGE defense network, and (2) linear programming in business and industry. This trend toward larger system problems will continue.

While business is constantly using more mathematics, the development of mathematical expressions for business problems is difficult and awaits further application of operations research and management science. Once mathematical models have been developed, then the establishment of criteria for optimum operation can be studied.

## Conclusions

The concept of feedback control is most valuable to business in specifying the factors of importance. These are:

1. System studies are needed to determine operation functions and the basic problems of business.

2. Business must be better able to measure performance.

3. Mathematical models are needed to describe business activities in order to evaluate alternate approaches, to study stability, and to forecast future events.

4. Optimizing criteria should be studied and evaluated.

It is likely that probabilistic theories will provide the most useful results. It is interesting to note that information theory, communication theory, and game theory are all probabilistic in nature.

# INFORMATION THEORY

by

E. M. Grabbe
Senior Staff Consultant on Automation
Computer Systems Division
The Ramo-Wooldridge Corporation

## Introduction

In any business, accounting and recording are directly concerned with information handling. A large part of the productive activities of a business is also concerned with information processing. These latter activities involve planning, design, drafting, scheduling, and control of machines. Hence, it is not surprising that management becomes excited when they hear that information theory may add new insight to, and perhaps solve, some of the knotty problems of data handling for business and industry.

Information theory is a new theory largely concerned with the quantitative measurement of information. While so far its results and applications have not been extensive, scientists in this country and throughout the world feel that it holds great promise for the future. It is generally considered that information theory is the greatest scientific advance in the past decade. Its application to problems at the management level and to business data processing problems will be in the future when the theory has matured and developed in more directions. However, it is worthwhile to examine the concepts the theory has developed so that we may understand more of its potentialities.

## Definitions

Information theory in the strict sense has been largely concerned with the quantitative measurement of information. It may be considered as a subdivision of a broader field called the statistical theory of communication which includes all the probabilistic analysis of communication. Some people use the term information theory in a broad sense to include the analysis of any situation dealing with the acquiring, transmittal, or use of information.

Information is the basis of all man's activities and provides the means of communication and direction between humans, between humans and machines, and between machines. There are two types of information that can be readily distinguished. One is pure scientific information and the other is what is called selective information.

Scientific information is concerned with description of fact and it is a more or less rigid body of data. For example, the formula $H_2O$ represents information about the structure of water.

On the other hand, selective information concerns events which have a certain probability of occurring. Information theory is concerned mainly with this type of information, its coding, and transmission.

## Measure of Information

The unit of information, the bit, is defined as that which makes a decision between two equally probable events. A very probable event has very little information associated with it whereas an improbable event has a higher level of information. In the case of flipping a coin we might assign the value 1 to a head and 0 to a tail. Hence, one bit of information is associated with the event of flipping a coin. We can see also that there is a probability associated with the results -- sometimes a head falls, sometimes a tail, hence this is selective information. With a trick coin having two heads, the information drops to zero since the results are clearly predictable. With a loaded coin one might flip three heads to every tail. Then more than one bit of information is required to describe the probability. Indeed two bits of information are needed since a tail has one chance in four of occurring.

A pair of dice have 36 possible combinations, representing sums from 2 to 12. The complexity of information is greater and more bits are required to describe an event. If a train runs on schedule day after day, this event has little information associated with it. However, if a train is delayed by one hour or so, any number of occurrences could account for the delay and the information required to describe what happened may be large.

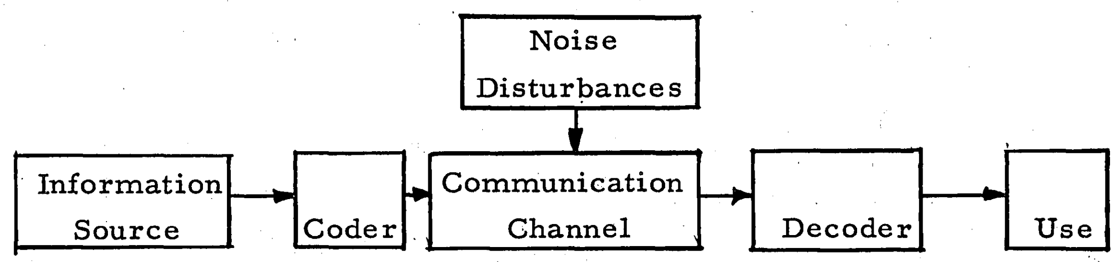
Consider the 1956 Series between the Yankees and Dodgers. Since the teams were fairly evenly matched, there was roughly an equal probability that either team would win the series. It would take but one bit of information to describe the winner. A "1" might mean that the Yankees won and a "0" that the Dodgers won. This is a very elementary amount of knowledge for it does not tell us how many games were played, what the scores were, in what innings the runs were scored, who the pitchers and catchers were, etc. It becomes necessary to have a large number of bits of information to distinguish between the various possibilities.

The manner of desdribing information in information theory implies that the exceptional case requires more information to describe than the routine matter. This is recognized in the newspapers in that "dog bites man" is not news but if "man bites dog" the information content is high enough to make the front page.

The measure of information is defined as $-\log_2 p$ where p is the probability of a given event.

## Communication System

A communication system to which information theory refers is shown below:




The source may be (1) a true generator of information -- a story teller making up a tale, (2) the output of a measuring instrument, or (3) a storage point for information generated elsewhere -- a stack of telegrams for transmittal or an output reel of magnetic tape from a computer.

A code consists of a set of symbols assigned to represent the information. This code is then used to transmit the information over a communication channel which may have disturbances called noise. Examples are radio and TV static, or cross talk on telephone lines.

To use the information it must be decoded. Information theory in its present state has not tackled the problems of the meaning and use of information.

## Coding

Once having this means of measuring quantity of information, we can then proceed to determine efficiencies of coding. For example, the Morse Code is used to transmit information, and four bits of information (dots or dashes) are required for the 26 letters of the alphabet. Morse recognized that certain letters in the alphabet occur very frequently and assigned them the shortest codes. For example, "e" is simply a dot. On the basis of information theory the Morse Code has 5.55 bits per symbol. Another code is the language that we speak. The coding of our English language is such that each character contains 2.2 bits of information. Discussion of language is largely theoretical since it is not likely that we will modify our language to improve its efficiency.

Other examples of codes are the catalog numbers of a mail order house or part numbers in an inventory. Decoding is the inverse operation of determining the description corresponding to the code.

## Results of Information Theory

Today we are engaged in large-scale transmission of information by radio, telephone, telegraph, and television. Improvements in efficiency of each transmission by improving the coding can be very worthwhile. Information theory gives us a relation between "bandwidth" and power and noise. It provides a method for finding the best possible code for a specific use. Finally, it predicts that ideally there is a way a communication channel no matter how noisy, can transmit information with negligibly small error. This final conclusion is astounding -- but, as yet, we do not know how to accomplish this in practical cases.

Information theory tells us the maximum rates at which transmission is possible over channels of given capacities, but again it does not indicate how one can approach this ideal limit. We know, for example, that in both radio and television we do not use the full capacity of the channels. To transmit a spoken message by telephone (40 bits of information per second), a channel capable of carrying 28,000 bits of information per second is used. Experimental devices can compress speech so that the required channel carrying capacity is reduced to 1/10.

A television picture requires about 200,000 bits of information to form a picture. Since an average word has five characters and it takes four bits of information per character, this is equivalent to 10,000 words. Hence, we see that the ancient Chinese were right in saying "a picture is worth 10,000 words".

## Man's Information Handling Capacity

It is interesting to compare man's ability to transmit or process information with a communication line. One can determine man's information handling capacity by the rate at which he can read information and reproduce it -- the rate at which shorthand may be taken or typing rates. In all of these cases man acts as a transmission line. It has been estimated that if man worked his full lifetime at transmitting information he could handle $50 \times 10^9$ bits of information (50 bits/second, 12 hrs/day, 60 years). This is equivalent to the information handled by a TV set in 16 minutes.

## Conclusion

It is clear that many of the concepts that appear in information theory are also associated with business operations. We may list these as follows:

1. Business operations are probabilistic in nature.
2. Coding of information is convenient and widely used.
3. Communication is an essential part of business operations.
4. Business wishes to minimize errors in transmission and handling of information.

Hence, it is likely that improvements in business data processing will come about through application of information theory. Such advances will first come through improvements in methods of coding and handling information in communication channels and in electronic computers.

# BOOLEAN ALGEBRA IN BUSINESS OPERATIONS

by

Eldred C. Nelson
Associate Director
Computer Systems Division
The Ramo-Wooldridge Corporation

## Introduction

Boolean algebra is a mathematical technique that appears to have promise for application to business operational problems. Although its actual use in solving these problems has, as yet, been very limited, its success in the solution of logical problems of engineering and science has suggested that it might also be useful in business.

Boolean algebra has some simple mathematical properties that are easily presented. Therefore, it seems appropriate to develop them first and then to discuss the applications with the aid of a specific knowledge of Boolean algebra.

## Boolean Algebra[1]

Boolean algebra is an algebra for dealing with things that have only two values, such as the on and off positions of a switch, the "go" and "no-go" of an inspection station, or the true or false property of a statement. These two values are represented mathematically by the two digital values:

$$0, \ 1.$$

---

[1] For a brief but straightforward exposition of the principles of Boolean algebra, see G. Birkhoff and S. MacLane, "Modern Algebra", The MacMillan Company, New York, N. Y., pp. 311-323; 1941.

These digits, 0 and 1, are the only numbers in Boolean algebra. There are no numbers 2, 3, 4, etc. All numbers in Boolean algebra are only one digit long[2] and they have either the value 0 or 1.

Boolean algebra is a method of manipulating the digits 0 and 1 in a convenient way. It can be expressed in terms of some elementary mathematical operations on the digits 0 and 1. An important elementary operation is the "complement" in which the digit 0 is exchanged for the digit 1 and the digit 1 for the digit 0. It is designated by a bar ¯ placed over the digit to be complemented. The nature of this operation may be represented by the following equations (the symbol "=" meaning equality in both directions):

$$\overline{0} = 1$$

$$\overline{1} = 0.$$

Other important elementary operations combine two digits to produce a third. One of them is called "logical addition" and is designated by the symbol "+" placed between the two digits. Its properties are shown in the following table:

$$0 + 0 = 0$$

$$0 + 1 = 1$$

$$1 + 0 = 1$$

$$1 + 1 = 1$$

The first three equations in this table are the same as those for addition in ordinary arithmetic. The fourth, however, is different. In ordinary arithmetic, the sum of 1 and 1 is 2, but there is no 2 in Boolean algebra;

---

[2] Although the binary digits 0 and 1 are used, Boolean algebra is quite different from the binary number system and should not be confused with it.

consequently, the right hand side of the fourth equation must be 0 or 1. The operation + is defined with the value 1 for the fourth equation.[3]

A third important elementary operation is called "logical multiplication". It is designated by the symbol ". " placed between the two digits. Its properties are shown in the following table:

$$0 \cdot 0 = 0$$

$$0 \cdot 1 = 0$$

$$1 \cdot 0 = 0$$

$$1 \cdot 1 = 1$$

These four equations are the same as those for multiplication in ordinary arithmetic. The three operations of complementation, logical addition, and logical multiplication are complete in the sense that all possible operations in Boolean algebra can be expressed in terms of them.

In addition to its arithmetic properties, Boolean algebra has the capability of treating symbolically the quantities that have only two values. Thus in the algebra, letter symbols, such as A, B, m, x, etc., called "variables", are defined to represent these quantities. These letter symbols may take on only the values 0 and 1. The mathematical operations described in the preceding paragraphs may be applied to the symbols. Thus, the equation

$$A + B = C$$

means that the value of A when combined with the value of B by the operation of logical addition must equal the value of C.

The three operations of +, ∘, and $\bar{}$ when applied to a single variable produce the following equations, which are satisfied identically for all values of the variable.

---

[3] If the value 0 were used on the right hand side of the fourth equation with the other three equations remaining the same, another operation of Boolean algebra would be defined. This operation is usually represented symbolically by $\oplus$.

$$A + \overline{A} = 1$$

$$A + A = A$$

$$A \cdot \overline{A} = 0$$

$$A \cdot A = A$$

(These equations are easily verified by substituting numerical values for A.)

Complementation has the following interesting property:

$$\overline{A + B} = \overline{A} \cdot \overline{B}$$

$$\overline{A \cdot B} = \overline{A} + \overline{B}$$

It is important to note that the operations of subtraction and division which are inverse to addition and multiplication in ordinary arithmetic, do not exist in Boolean algebra. Thus from

$$A \cdot B = A \cdot C$$

it does not follow that

$$B = C.$$

Similarly, from

$$A + B = A + C$$

it does not follow that

$$B = C.$$

The concept of a "function" is useful in Boolean algebra. Its definition is the same as in ordinary mathematics: A function of a variable (or variables), say A, is a rule by which there is assigned to each value of that variable (or set of values of those variables) a function value F (A). Function values may only have the values 0 or 1 too.

Owing to the restriction to the digits 0 and 1, there are only four possible functions of one variable. They are:

| A | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 |

In algebraic symbolic form, these functions may be written:

$$F_1 = 0$$

$$F_2 = A$$

$$F_3 = \overline{A}$$

$$F_4 = 1$$

There are sixteen possible functions of two variables. In general, there are $2^{2^n}$ functions of n variables.

## Applications in Engineering and Science

Boolean algebra has been applied extensively in the design of the switching networks of digital computers[4] and of telephone systems. The elementary switches in this equipment have only two operational states, "off" and "on"; consequently, they can be represented mathematically by binary digits 0 and 1 and the designing of networks of these switches can be treated by Boolean algebra. It is instructive to consider an example from the design of switching networks. Let 0 represent the off state of a switch and 1 represent the on state and let the switching properties of various switches be represented by letter symbols A, B, C, D, etc.;

---

[4]For an exposition of the use of Boolean algebra in digital computer design, see E. C. Nelson, "An Algebraic Theory for Use in Digital Computer Design", Proc. Trans. IRE PGEC EC-3, 12, (1954).

i.e., the symbol A represents the on-off state of one switch, the symbol B represents the on-off state of another switch, etc. If A and B are two switches, another switch C can be made by connecting A and B as specified by the equation

$$C = A + B.$$

This equation states that C is a switch that is made by connecting together the switches A and B so that the resultant switch C is on if either A or B or both are on and is off only if both A and B are off. Similarly, the equation

$$D = A \cdot B$$

states that D is a switch that is made by connecting together the switches A and B so that the resultant switch D is on only if the switches A and B are both on; otherwise it is off.

Boolean algebra was developed initially by George Boole for symbolic logic.[5] In logic, an important elementary concept is the property of a statement being true or false. Since true and false are two values, they can be represented by the binary digits; usually 0 represents the statement being false and 1 represents the statement being true. Letter symbols also are used. Thus the statement

A dog is an animal

is either true or false. This property of the statement may be represented by a letter symbol, say A. A having the value 0 means that the statement is false and A having the value 1 means that the statement is true. In symbolic logic, each statement is assigned a letter symbol which represents its truth property. When statements are combined to form new statements, their truth symbols can be combined in a similar manner so that the resultant mathematical expression represents the truth property

---

[5] For a readable exposition of the principles of symbolic logic see Alfred Tarski, "Introduction to Logic and to the Methodology of Deductive Sciences", Oxford University Press, New York, N.Y; 1946.

of the new statement; e.g., the statement

A fish is an animal

having a truth property represented by the symbol B may be combined with the preceding statement to produce a new statement

A dog is an animal and a fish is an animal

having a truth property C represented by

$$C = A \cdot B.$$

As statements in the English language, the first one is true, so A has the value 1, and the second one is false, so B has the value 0. The statement resulting from combining the two statements with the conjunction "and" is false, so C should have the value 0, which is just the value calculated from the equation for C in terms of A and B.

On the other hand, the statement obtained by combining those statements with the conjunction "or"

A dog is an animal or a fish is an animal

is a true statement and its truth property, say D, should have the value 1. D may be expressed in terms of A and B by the equation

$$D = A \dotplus B.$$

Substitution of the values for A and B into this equation furnishes a value of 1 for D.

Thus in symbolic logic, statements can be combined by means of conjunctions and the truth property of the resultant statement can be calculated by applying Boolean algebra to the truth properties of the statements.

## Application to Business Operations

Boolean algebra can be used to treat mathematically those problems in business operations in which the elementary things have only two values. The only applications in business operations known to the author are applications in symbolic logic in which the elementary things are the truth properties of elementary statements.

Symbolic logic has been applied to the analysis of certain types of business policy. A policy is usually a collection of statements about certain aspects of the business. To apply symbolic logic, these statements are analyzed into a set of elementary statements and combinations of the elementary statements. The truth properties of the elementary statements are represented by letter symbols and the truth properties of the combination statements by Boolean algebraic equations containing those symbols. Analysis of the policy by Boolean algebra can test the policy for consistency and completeness. Symbolic logic has been applied to the analysis of insurance contracts to test them for consistency and completeness.

Applications of Boolean algebra to business operational problems have been very limited. It is a mathematical technique that seems to have promise for application to these problems and therefore some interest in it has been aroused in the business community.

# NONLINEAR PROGRAMMING

by

Andrew Vazsonyi
Head, Management Sciences Department
Computer Systems Division
The Ramo-Wooldridge Corporation

A certain corporation is involved in producing and marketing three different products, all utilizing the same type of resource. By the word "resource" we might mean labor resources or the land required in a farm enterprise or warehouse space, or any other commodity that goes into the production of these three products. Figure 1 shows the profit for each of these products, and it is assumed that any quantity of these products can be sold. For instance, the top figure in Figure 1 shows that if product 1 is produced in quantities of less than 15 units, then the profit is 1.2 million for each unit produced. (The million is omitted from the graph in order to save space.) However, if more than 15 of these units are produced, then the profit drops down to $100,000 for each product. The other graphs in Figure 1 have similar meaning. The problem we propose to examine now is the following: the total resources available to the corporation are given--how to allocate these resources to the three products?

Let us denote by A the total resources available, and let us denote by $x_1$, $x_2$, and $x_3$ the quantities of these resources allocated to the first, second, and third product. Then

$$x_1 + x_2 + x_3 = A \qquad (1)$$

The profit associated with this allocation of products can be written as

$$z = f_1(x_1) + f_2(x_2) + f_3(x_3) \qquad (2)$$

where each of these particular functions refer to the graphs shown in Figure 1. Mathematically speaking, then, the problem is to maximize the profit as expressed by Equation (4) under the restriction of Equation (39). This problem is often referred to as the problem of the distribution of effort as the resources are considered as efforts expended in producing these different products. The problem we are describing here is a problem in mathematical programming, but it is not a linear programming problem. The reason is that the payoff function as expressed by Equation (2) is a nonlinear function. In simpler language, we can say that the functions shown in Figure 1 are not straight lines, whereas in linear programming problems, the payoff function must be straight.

In the example we have described, the payoff function was nonlinear but the relation between the unknowns was linear. If we replaced (1) by

$$x_1^2 + \log x_2 + x_3 = A \tag{3}$$

we would have a nonlinear relationship between the unknowns. It is seen, then, in a linear programming problem, all the equations must be linear; in a nonlinear programming problem some or all the equations may be nonlinear.

In the case of linear programming problems there is a general method, the simplex method, which allows the computation of the solution which minimizes (or maximizes) the payoff function. In the case of nonlinear problems no such general method is available. Some nonlinear problems can be solved, some cannot. Each problem must be examined on its own.

To illustrate our point, we will show now how the allocation problem we described here can be considered as a problem in "convex" programming and how it can be solved.

The functions shown in Figure 1 are called by mathematicians, convex from below. The curve shown by the solid line in Figure 2, or the dotted polygon shown in Figure 2 are both convex from below. The definition of convexity is further explained with the aid of Figure 2. Namely, if one chooses any two points on the graph, say A and B, and one selects any points, say C, situated on the straight line connecting A and B, then C <u>lies below the graph</u>. Figure 3 shows a graph which is not convex: D lies above the graph.

The theory of convex programming is based on the following theorem: if in a mathematical programming problem all the equations relating the unknowns are linear and the payoff function is convex from below, then the payoff function can be minimized with the aid of the techniques of linear programming.

As an illustration of this theorem, we show now how the allocation problem discussed above can be formulated as a problem in linear programming.

We denoted by $x_1$ the resources allocated to the first product. Let us denote by $y_1$ that fraction of the resources allocated to the first product for which the unit profit is 1.2 million. As there cannot be more than 15 units that yield this rate of profit, we have

$$0 \leq y_1 \leq 15. \quad (4)$$

Now, we denote by $y_2$ the fraction which yields the unit profit of $100,000. Here

$$0 \leq y_2 \leq 20 \quad (5)$$

(It can be observed on the top graph in Figure 1 that there cannot be sold more than 20 units of this type.)

Quite similarly, we divide $x_2$, the resources allocated to the second product, into three parts, $y_3$, $y_4$, $y_5$, as suggested in the middle graph in Figure 1. Here we have

$$0 \leq y_3 \leq 20 \quad (6)$$

$$0 \leq y_4 \leq 20 \quad (7)$$

$$0 \leq y_5 \leq 10 \quad (8)$$

Finally, we divide $x_3$ into three parts, $y_6$, $y_7$ and $y_8$, as suggested by the bottom graph in Figure 1. We have here

$$0 \leq y_6 \leq 10 \quad (9)$$

$$0 \leq y_7 \leq 5 \quad (10)$$

$$0 \leq y_8 \leq 15 \quad (11)$$

Let us remember now that

$$y_1 + y_2 = x_1 \quad (12)$$

$$y_3 + y_4 + y_5 = x_2 \quad (13)$$

$$y_6 + y_7 + y_8 = x_3 \quad (14)$$

Now we observe the following important point. The profit can be written as

$$z = 1.2y_1 + .1y_2 + .8y_3 + .4y_4 + .3y_5 + 1.5y_6 + 1.0y_7 + .6y_8 \quad (15)$$

and this is a linear payoff function. We see then that with the aid of the now variables, we can express our problem as a linear programming problem.

As we stated before, this is always the case for convex programming problems and, therefore, we see (at least for this special case), that convex programming problems can be solved with the aid of linear programming.

We see then that here we have a method which allows the solution of certain nonlinear problems. Now we proceed to discuss another type of nonlinear problem which can also be solved, though not with the aid of techniques of linear programming.

## A Problem Related to the Attrition of Production Facilities

The production department of a certain corporation is making long-range plans for the assignment of production facilities. Suppose that the corporation has 1000 special automatic machines which can produce only two different products. If x number of these machines are allocated to produce the first product, the profit will be represented by the upper f(x) curve shown in Figure 4. On the other hand, if y number of machines are assigned to the second product, the profits are shown by the lower curve g(y) in Figure 4. We wish to make a decision for each of the following four years with respect to the question of how many of each of the machines should be allocated to the first and second product. Observing that profits on the first product are higher than for the second product, one would think, at first, that all production should be allocated to the first product.

There is, however, another factor that must be considered: the attrition rates of the machines are different. Namely, if 100 machines are assigned to the first product, then after the first year there will be only 60 machines in working condition. On the other hand, if 100 machines are assigned to the second product then there will still be 80 machines in working condition at the end of the first year. Now at the end of the first year, that is, at the beginning of the second year, there will still be a certain number of machines available, depending on how the machines are allocated during the first year. We assume that these machines will be reconditioned and will be as good as new. Now, a new decision has to be made on how to allocate the machines in the second year of production. At the end of the second year of production a decision is to be made again, and finally at the end of the third year, a final decision is to be made as far as the allocation of machines for the fourth year is concerned.

In order to make the problem more manageable, we assume at the end of the fourth year that the machines will all become obsolete and they will have no value. It can be seen, then, that this factor of attrition makes the problem more complicated, and we propose now to proceed to set up the mathematical equation for the solution of this problem.

Let us assume that we allocate $x_1$, $x_2$, $x_3$, and $x_4$ number of machines to the first product during each of the four years, and we allocate $y_1$, $y_2$, $y_3$, and $y_4$ number of machines to the second product. We assume that we start with $A_1$ number of machines, and so we have

$$x_1 + y_1 = A_1. \tag{16}$$

(In our particular case, the value of $A_1$ is 1000.) Let us denote by $A_2$ the number of machines in working condition at the beginning of the second period. Then we have

$$A_2 = .6x_1 + .8y_1. \tag{17}$$

Consequently, we must have

$$x_2 + y_2 = A_2. \tag{18}$$

Quite similarly, we get for the third and fourth year, the following equations:

$$A_3 = .6x_2 + .8y_2 \tag{19}$$

$$x_3 + y_3 = A_3. \tag{20}$$

and

$$A_4 = .6x_3 + .8y_3 \tag{21}$$

$$x_4 + y_4 = A_4. \tag{22}$$

The total profit for the four years can be expressed by

$$z = f(x_1) + f(x_2) + f(x_3) + f(x_4) + f(y_1) + g(y_2) + g(y_3) + g(y_4). \tag{23}$$

It can be seen then that mathematically speaking the problem is to maximize the profit as given by Equation (23) under the restrictions of the Equations (16) to (22). This problem is not a problem in linear programming, as the profit is not a straight line or linear function. In order to solve such a problem, the techniques of linear programming are not adequate. However, we will show that a method of solution for this nonlinear program can be developed.

Let us first plan for a single year and assume that a total of A machines are available at the beginning of this year. (It is not assumed that A has the value of 100.) Let us assume that x of these production machines are allocated to the first product, and y of the machines are allocated to the second product. The profit associated with this method of allocation is given by

$$z = f(x) + g(y) \tag{24}$$

where

$$x + y = A. \tag{25}$$

It is recognized, then, that if we plan only for one year we have the problem of the distribution of effort. In order to solve this problem, we first prepare a table of numbers where for different values of x and y we enter the appropriate value of the profit z. Then we select a numerical value of A, survey the profit values on the table for which x+y=A and select the maximum profit that can be obtained for this numerical value of A. We can repeat this procedure for different values of A and thereby obtain the largest profit that can be obtained with A machines, provided we are planning ahead only for a single year. Let us denote by $F_1(A)$ this largest profit.

This method of computing the table and selecting the maximum profit is equivalent to the following mathematical statement:

$$F_1(A) = \max_{0 \leq x \leq A} \left[ f(x) + g(A - x) \right] . \tag{26}$$

Let us consider now a two-stage problem, and assume again that we begin with A number of machines, and that we allocate x of these to the first product and y of these to the second product. According to our attrition rate, the number of machines available for the second period is given by

$$A' = .6x + .8y. \tag{27}$$

Now, whatever way we allocate our machines in the first period, we certainly will allocate our machines in the best possible fashion in the second period. Therefore, we can say that our profit for the two years of the planning plan is given by

$$z = f(x) + g(y) + F_1(.6x + .8y). \tag{28}$$

However, we recognize that y is given by

$$y = A - x \tag{29}$$

and the number of machines available at the beginning of the second year is given by

$$A' = .6x + .8y = .6x + .8(A - x) = .8A - .2x. \tag{30}$$

Therefore, we can say that the combined profit for the two years is given by

$$z = f(x) + g(A - x) + F_1(.8A - .2x). \tag{31}$$

The question is, then, what value of x should we select to make this profit the largest possible. Now by assuming various values of A and x, we can prepare a table showing the variations in the above costs. Then for each value of A we can select the value of x which gives the maximum profit. In mathematical notation we say that the best profit is given by

$$F_2(A) = \max_{0 \leq x \leq A} \left[ f(x) + g(A - x) + F_1(.8A - .2x) \right], \tag{32}$$

provided that we start with A machines and consider only a two-year planning period. Quite similarly, if we denote by $F_3(A)$ the best possible profit that can be obtained in a three-year period, we can determine this $F_3(A)$ with the aid of the equation

$$F_3(A) = \max_{0 \leq x \leq A} \left[ f(x) + g(A - x) + F_2(.8A - .2x) \right]. \tag{33}$$

This is again the mathematical statement describing the preparation of a table of numbers and the selection of the appropriate maximum values.

We can proceed now to solve the problem for the four-year period, and therefore it can be seen that the whole problem can be solved by a step-by-step process. Each step involves the preparation of a table of numbers--the selection of the highest payoff possibilities from this table and each step relies on the result of the previous step. Such a procedure is referred to in the literature as dynamic programming, and there are many mathematical programming problems that can be solved by the methods of dynamic programming.

## Other Nonlinear Programming Problems

As we have already stated, there are no general methods for solving nonlinear programming problems and each case must be solved on its own. Some nonlinear programming problems can be solved with the aid of calculus, and here the method of using Lagrangian multiples is an important tool. There are some nonlinear problems which can be solved by graphical methods, while there are some which can be solved by using electronic computers. Generally speaking, however, each business problem that leads to a problem in nonlinear programming, presents a new challenge to the investigator.

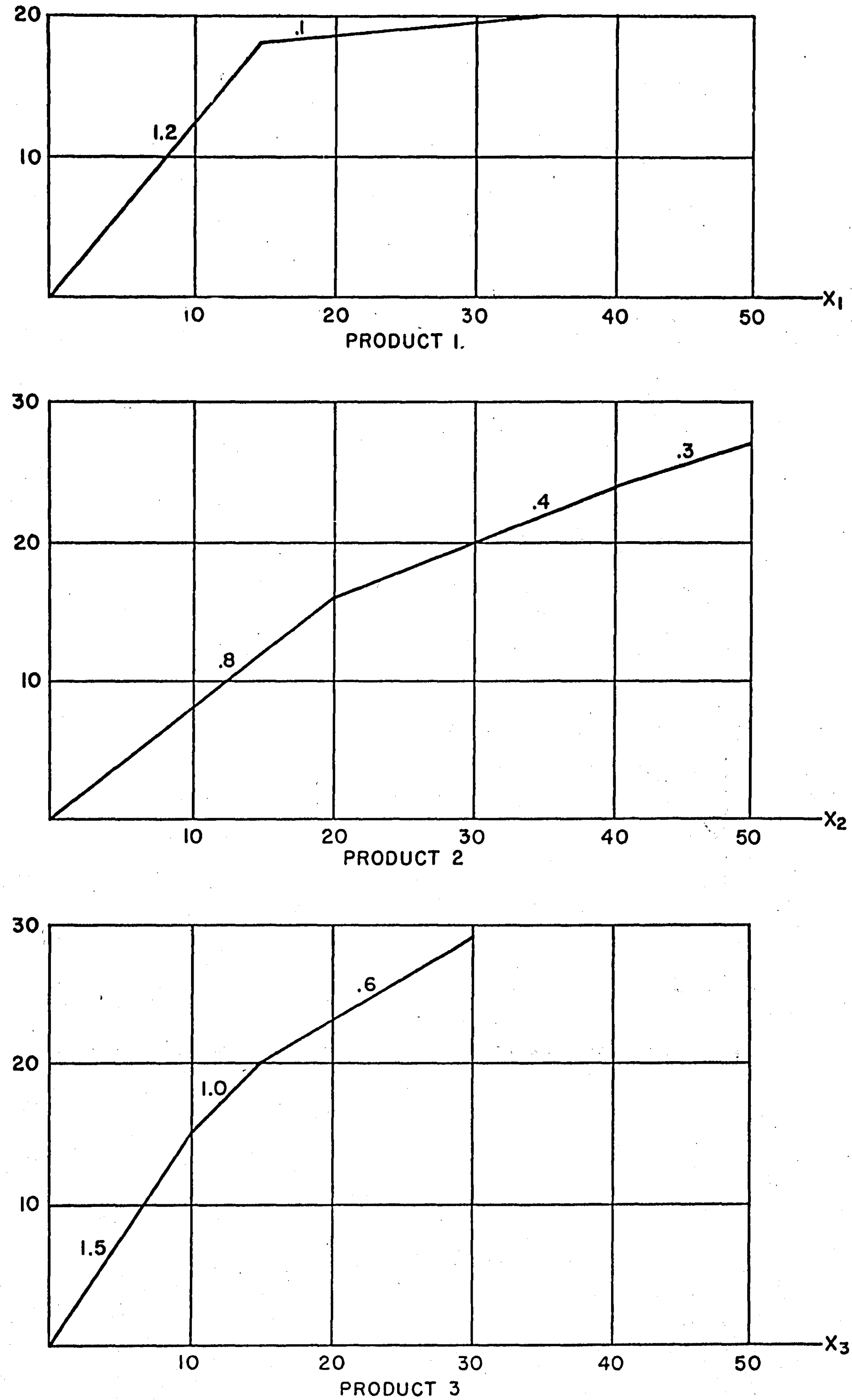



FIGURE I. PROFITABILITY OF EACH PRODUCT

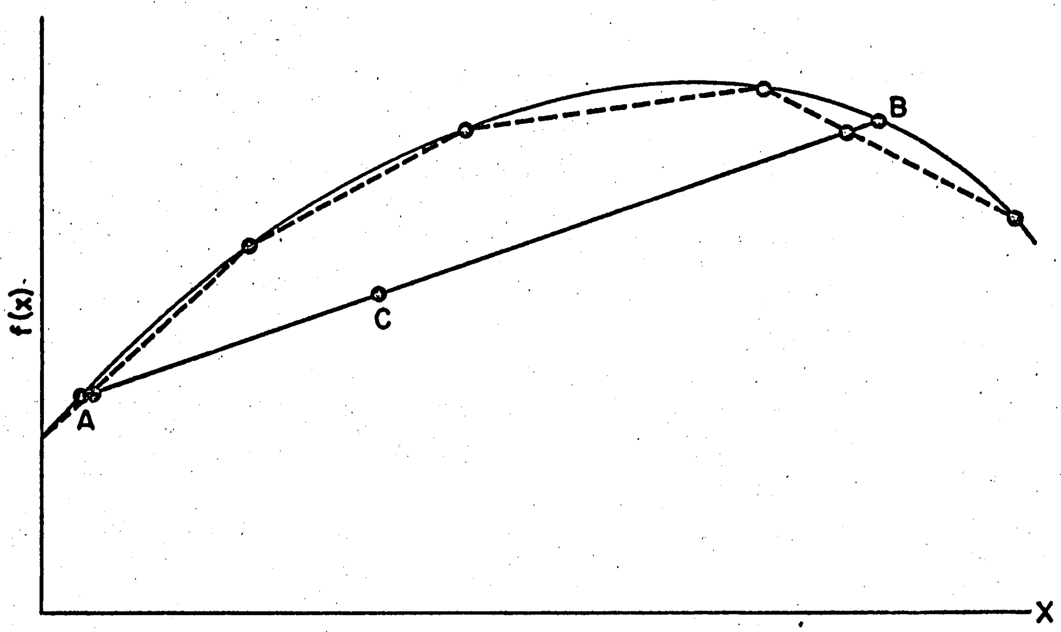



FIGURE 2. THIS CURVE IS CONVEX FROM BELOW. POINT C IS BELOW THE CURVE.

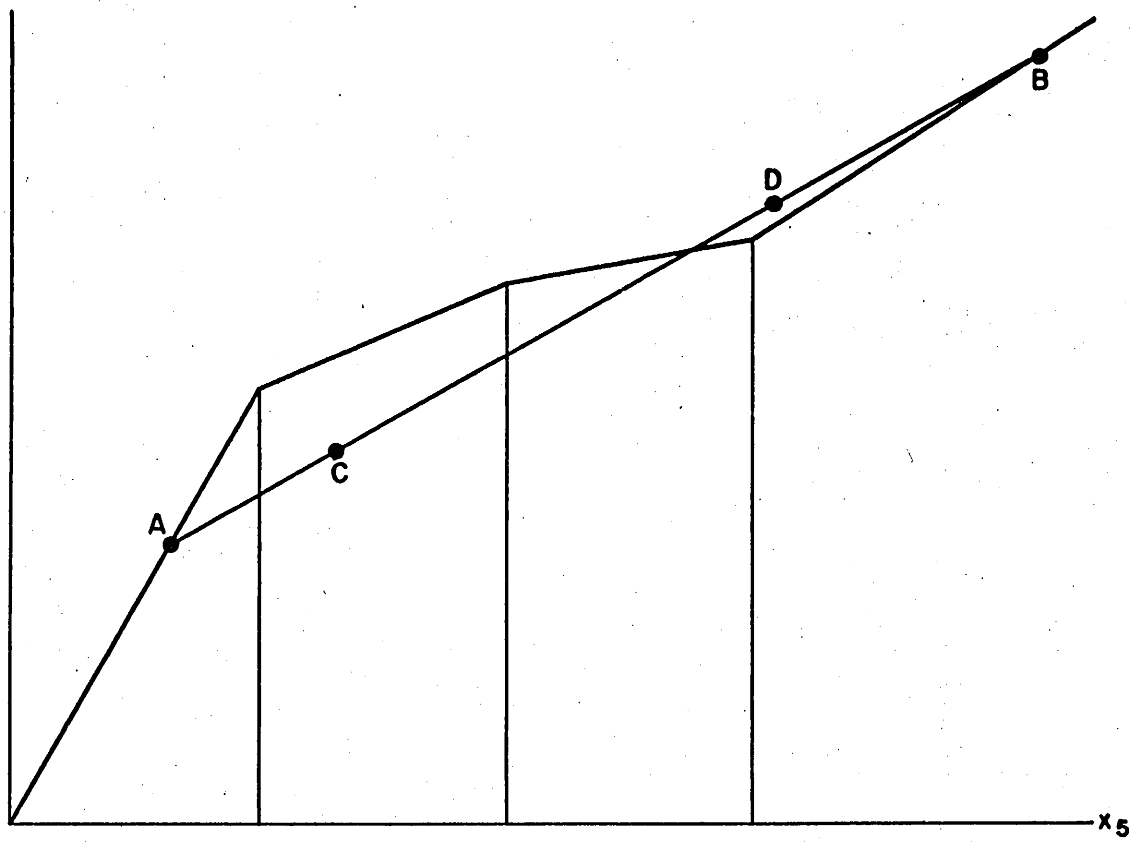
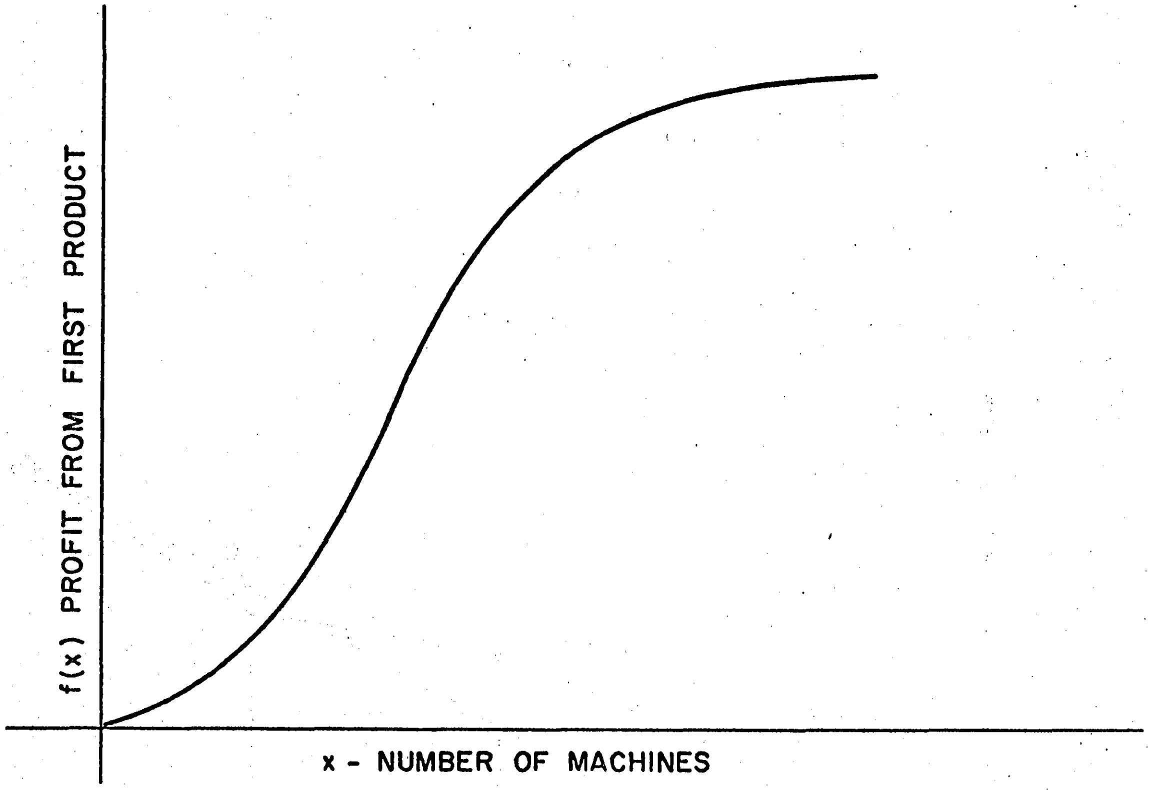



FIGURE 3. NON CONVEX PAYOFF FUNCTION. POINT C IS BELOW, POINT D IS ABOVE THE CURVE.



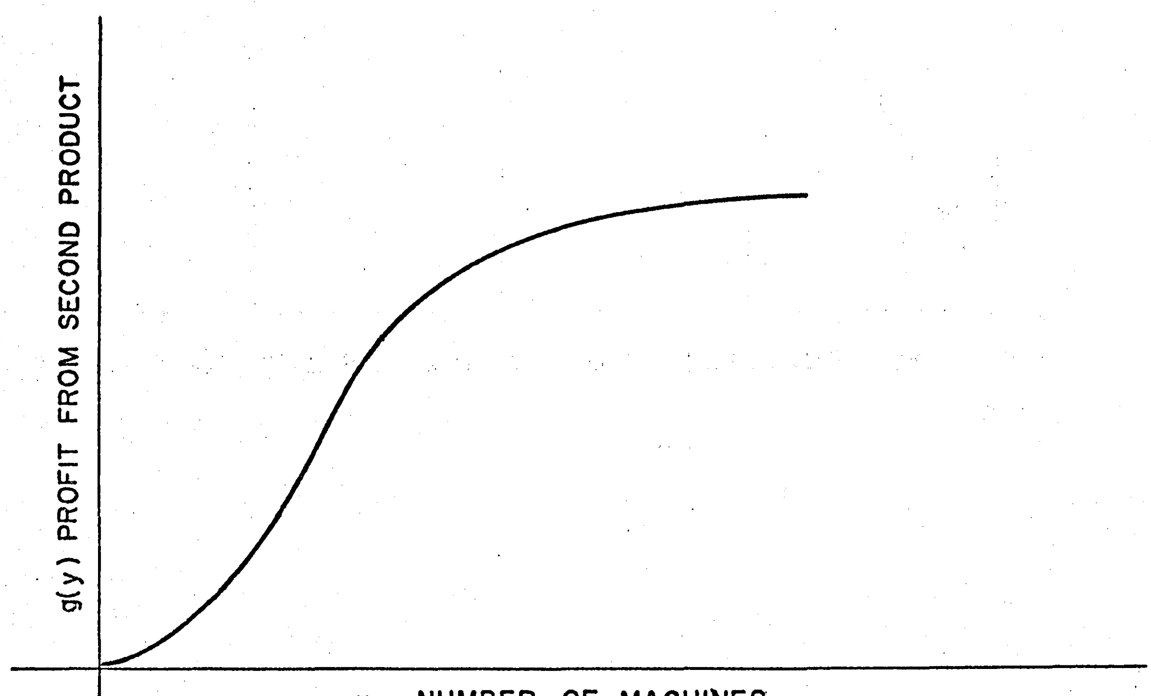




FIGURE 4. A PROBLEM RELATED TO THE ATTRITION OF PRODUCTION FACILITIES

# THEORY OF GAMES

by

Andrew Vazsonyi
Head, Management Sciences Department
Computer Systems Division
The Ramo-Wooldridge Corporation

Suppose I am the manager of a refinery, and I am producing insufficient quantities of gasoline and insufficient quantities of Diesel oil, but I produce a surplus of Bunker fuel oil. This condition makes it necessary for me to export the surplus of Bunker fuel and import quantities of gasoline and Diesel oil to meet my market requirement. This, of course, leads to excess transportation costs and lowers my profit. My staff recommends four different methods of improving my profit structure. These proposals recommend certain policies; namely, that I modernize some of my production facilities and that I introduce some new promotional means to market my product. Each of these proposed policies leads to a different increase in profit. My current profits are 30 cents per barrel. If I accept the first recommendation, profits go up to 35 cents, the second recommendation leads to 39 cents, the third one to 42 cents, and the fourth one leads to 44 cents per barrel. (These figures include depreciation costs associated with the capital expense required to modernize my refinery.) Which of these recommendations, if any, should I accept? At first glance, it would appear that obviously I should take the fourth proposal which leads to a profit of 44 cents per barrel.

A more complete examination of the facts, however, uncovers some additional important information.* The total market in this area is essentially fixed and split between my refinery operation and my principal competitor--there is no way of changing the total market requirement. Furthermore, I know that my competitor is considering changing his refinery equipment and also, considering other ways of improving his profit. As I know a great deal about my competitor's operation, and I know the market, I can figure out the possible improvements he is considering. Evaluation of his possible actions lead to a table of potential profits per barrel of product, as shown in Figure 1.

---

*Symonds, G. H., Application of Linear Programming to the Solution of Refinery Problems, Esso Standard Oil Company, Linden, New Jersey

This figure shows that if my competitor does not make improvements, and I make no improvements, I receive 30 cents profit per barrel. Making no improvements is designated as my first policy, or policy number (1) . If I accept the first recommendation of my staff, that is, policy (2) , and he makes no improvements, the figure shows that I get 35 cents per barrel; according to policy (3) 39 cents; according to policy (4) 42 cents; and according to policy (5) 44 cents. We designate his policy of making no improvements as his first policy, or policy number [1] . However, if he proceeds with policy [2] as a method of improving his business, my profit will change. If my policy remains unchanged, I get 34 cents a barrel; if I introduce my policy (2), I get 38 cents; my policy (3) leads to 41 cents, etc. Figure 1, then, shows my profit, whatever the policy accepted by my competitor, and whatever the policy accepted by my own refinery.

What should my action be in this case? Suppose I accept my policy (5), what would happen? If he doesn't change his policy, I get 44 cents per barrel. On the other hand, if he "plays" policy [5], I only get 30 cents per barrel. Perhaps I should accept my policy (3) ? Then I would get at least 37 cents profit on each barrel; or if he does not select his best policy (which is [5] ), I might even get more than 37 cents per barrel. However, the question is this, is my policy of accepting policy (3) the best possible? Is it possible for me to assure a profit of more than 37 cents per barrel?

It is clear from Figure 1 that if I choose a policy and adhere to it, then policy (3) is the best. However, there is another possibility; namely, that I change my policies from month to month, and in this way perhaps I can improve my profit to a higher figure than 37 cents per barrel. In fact, perhaps I could prepare a program for the next year, specifying in what month which of the policies I am going to "play". However, I do not consider this plan practical, because it is very likely that my competitor would find out about my plan and then he could counteract it by selecting an appropriate sequence of policies.

At this point, I hit on the following plan: suppose that at the beginning of each month I toss a coin. When I get heads, I "play" policy (2), when it is tails, I "play" policy (5). In this way at least I can be sure that my competitor is not going to know what policy I am going to "play". This "game" is somewhat unsatisfactory because I can select from only two different policies. Suppose I throw some dice and according to the numbers that come up, I select my policy. Another possibility would be to take a big urn and fill it with tokens which have the numbers 1, 2, 3, 4, and 5 written on them. I pull one of these tokens out and the number shown on its face is the policy to be played that particular month. The question now is, how many tokens should have the number "1"? How many should have the number "2", etc. Is it possible by playing this game to get more than 37 cents profit on each barrel of product? If the answer is affirmative, I would like to know what would be the best possible way of marking the tokens in order to get the highest possible profit, and what is the value of this highest possible profit.

The problem we have formulated here is a problem in the branch of mathematics called "Theory of Games". This mathematical theory leads to the answer to the problems posed here--place a single token into the urn with the number "2" written on it, and put in two tokens with the number "3" written on them. Pull one of the three tokens from the urn and play the particular policy that is written on the token. Replace the token and pull one out again next month, etc. On the basis of "Theory of Games" we can be certain that this is the best possible policy that we can adopt; this policy leads to an average profit of 37 2/3 cents per barrel. In other words, using this scheme improves the profit by 2/3 of a cent per barrel.

To be more specific, accepting this policy insures getting this 37 2/3 cents per barrel of profit. Furthermore, on the basis of the "Theory of Games" I also know that if my competitor "plays" the same sort of "game", he should place into his urn one token with his policy [1] and two tokens with his policy [5]. Then my competitor will restrict me to the long-run profit of 37 2/3 cents per barrel. On the other hand, if he "plays" any other "strategy" my profit will be higher than this 37 2/3 cents.

The game we describe here is called by mathematicians a zero-sum two-person game. The table of profits shown in Figure 1 is called the payoff matrix of the game. We have here a five-by-five payoff matrix, since we have five rows and five columns in the table.

## A Two-By-Three Two-Person Game

In order to develop the basic concepts of the Theory of Games, we will examine in detail a two-person game which is somewhat simpler than the one we have considered so far. We assume that there are only two strategies available to me, and there are three strategies available to my opponent and that the payoff matrix is given by

| | [1] | [2] | [3] |
|---|---|---|---|
| (1) | 2 | 3 | 11 |
| (2) | 7 | 5 | 2 |

The question is, what strategy should I play?

I might take a very optimistic point of view and say that I am after the largest possible gain. This is the \$7 shown in the lower left-hand corner in the matrix. So, I say I am to play strategy ②. However, as time goes on, my opponent is going to find out that I play strategy ② and he is going to counteract my strategy by playing his strategy [3]. In this case, then, my gain is only \$2. Furthermore, suppose now that in some other game, there is a fourth strategy for the opponent, with payoffs of \$1 and \$6. Under this condition, my opponent would never play his strategy [1] as in strategy [1] he has to pay \$2 or \$7, while in this hypothetical fourth strategy he has to pay only \$1 and \$6. This means that his strategy [4] is always better than his strategy [1] (as he always loses less money by playing strategy [4]). Therefore, it would be in vain for me to hope for the \$7 gain, as he would never play strategy [1]. It is seen, then, that my policy of going after the largest possible payoff is not necessarily a rational one.

According to the argument we followed in the case of the case of the oil refinery, we suspect that I should play a combination of my two strategies, that is, I should alternate between strategy ① and strategy ②. Perhaps I should play ① half of the time, and strategy ② also half of the time. Suppose for the moment, that my opponent always plays his strategy [1]. How much money am I going to get in the long run, if I play 50% of strategy ① and 50% of strategy ②. Clearly, the answer is halfway between the gain of \$2 and \$7, that is, my gain in the long run will be \$4.50.

We propose now to extend this argument to a somewhat more complicated case. Let us assume again that my opponent still plays his strategy [1], but that I play a fraction x times strategy ① and the rest of the times, that is, (1-x times), I play strategy ②. What is my gain under these conditions? In the long run, I will get x times \$2, and 1-x times \$7, which gives me

$$z = 3x + 5(1 - x) = 5 - 2x. \qquad (1)$$

A graphical representation of this equation is shown in Figure 2 by the line [1]. It can be verified that if x=.5, that is, if I play my strategy ① .5 times, then the gain is, indeed, \$4.50. If I never play my strategy ①, x is zero and the gain is \$7, as it can again be seen in the figure. Furthermore, if I play strategy ① all the time, then x is 1 and the gain is \$2. It is seen, then, that in Figure 2 the straight line [1] shows any possible gain I can get, as long as my opponent plays strategy [1].

Quite similarly, we can compute the possible gain if my opponent plays strategy [2]. This gain is given by

$$x = 3x + 5(1-x) = 5 - 2x \tag{2}$$

and is shown graphically in Figure 2 by the straight line marked with [2]. Finally, if my opponent plays strategy [3], then my gain is described by

$$z = 11x + 2(1 - x) = 2 + 9x \tag{3}$$

which is again shown in Figure 2 by the straight line [3].

Incidentally, if I play only strategy (1) or only (2), this is described in the theory as playing a <u>pure</u> strategy. We observe that in Figure 2, pure strategies are represented by either the value x=0 or x=1, and, therefore, the gains for pure strategies are given by the left-hand or right-hand side verticals.

We noticed that being a optimist is not necessarily a rational way of playing the game. Another possible method of argument is "the principle of insufficient reason". I take the point of view that I have no idea what my opponent is going to play and, therefore, I arbitrarily assume that one-third of the time he is going to play strategy [1], one-third of the time strategy [2], and one-third of the time strategy [3]. Again, let us assume that I play strategy (1) x fraction of the time. Now let us ask the question, what is my gain under these conditions? The way to compute this gain is to take the gain shown by Equation 1 multiplied by one-third, take the gain shown by Equation 2 multiplied by one-third, take the gain shown by Equation 3 multiplied by one-third, and add up these three gains. On the basis of the principle of insufficient reason, then, we get the following gain:

$$z = 1/3(7 - 5x) + 1/3(5 - 2x) + 1/3(2 + 9x). \tag{4}$$

This can be simplified to

$$z = 14/3 + 2/3x. \tag{5}$$

In Figure 2 this line is shown by the dotted straight line. For instance, it can be seen that if I play my strategy (1) half of the time and my strategy (2) half the time (and I assume that my opponent will play his three strategies in equal proportion), I can expect a gain of $5, (point P in the figure). In order to make my gain the largest, I should mix my strategies such that I reach the right-hand upper end of the dotted straight line. This means, that I should play my strategy (1), all the time, that is, I should have the pure strategy of (1). Then I can hope for a gain of 16/3 or $5.30, (point Q in Figure 2).

Let us now ask the question whether this principle of insufficient reason makes good sense. If I play my strategy (1) all the time, my opponent is going to find this out and then he is going to counteract by playing his strategy [1] and so I will collect only $2 instead of the $5.30 I had hoped for. We can see, then, that the principle of insufficient reason is not necessarily a rational argument for deciding what strategy I should use.

We propose now to examine another method of arguing, which one could describe as the conservative attitude. If I play the pure strategy (1), I am quite certain to get at least $2. (I get exactly this amount if my opponent plays his strategy [1].) If he plays some other strategy, or a combination of his other strategies, I am going to get more than $2. However, I am quite certain that if I play strategy (1), I will get at least $2. Suppose now I play my pure strategy (3). It happens to turn out (this is just a coincidence) that in this case again I can be certain of getting at least $2. Now, I ask the question, if I should mix strategies (1) (2) and (3) in some fashion, what is the number of dollars I can be certain of getting? If I play my strategy (1) x number of times, then the respective gains, depending on whether he plays strategy [1], [2], or [3] is given by one of the following three numbers:

$$7 - 5x \qquad 5 - 2x \qquad 2 - 9x \tag{6}$$

I am certain that I am to get at least as much as the smallest of these three numbers. To illustrate, let us assume that x=.4, that is, I play strategy (1) 40% of the time. If he plays his strategy [1], I am to get $5, as shown by point A in Figure 2. If he plays his strategy [2], I get $4.20 as shown by B, and if he plays his strategy [3] I get $5.60 as shown by point C However, I do not know which of these strategies he is going to play. All I know is that I can be certain of getting $4.20 which is the smallest of the three numbers. I can put this thought into mathematical form by saying that whatever number of times I play strategy (1), that is, for any value of x I am certain to get the gain given by

$$z = \min(7 - 5x,\ 5 - 2x,\ 2 + 9x). \tag{7}$$

The word ahead of the bracket is an abbreviation for minimum, and it means that one is to take the minimum (that is, the smallest) of the three numbers. Now, I ask, what value of x should I use? It will make good sense to select x, so that this gain that I am certain to get becomes the largest possible. The three straight lines in Figure 2 form a shape somewhat like a roof of a house, and our problem is to find the top of this roof, at which point my certain gain becomes the largest. In order to determine this point D on the diagram, we can either read the figures as shown, or we can compute this intersection by solving the equations for the two straight lines. The equations for these two lines are given by Equations (2) and (3), and by solving them we get the value x=3/11. This gives the fraction of times I should play my strategy ①. The remaining time, that is, 8/11 time, I should play strategy ②. It is easy to compute that my expected gain is 49/11 dollars, which is about $4.50. This gain is called the value of the game and we denote this gain by V. It is seen, then, that mathematically speaking, this gain is determined by selecting the value of x, which maximizes the gain as given by Equation (7). This means then that the value of the game for this particular game we are discussing is given by the formula

$$V = \max_{0 \leq x \leq 1} \min(7 - 5x,\ 5 - 2x,\ 2 + 9x) . \tag{8}$$

In summary, then, we see that if I play my strategy ①, 3/11 times, and my strategy ②, 8/11 times, then whatever my opponent does, I will still get $4.50. The conservative point of view leads to a consistent method of solving our problem.

## The Strategy of the Opponent

Let us now put ourselves into our opponent's shoes. What is the strategy that he should play? Again, we accept the concept of visualizing the worst, and so we will say that his problem is to be certain to prevent me from gaining too much money. In other words, he should try to keep my gain down. If he plays pure strategy [1], all he can be certain of is that he will keep me below $7. (See Figure 2.) This is due to the fact that if I play pure strategy ② and he plays his pure strategy [1], I get exactly $7. We suspect again that his strategy should be a mixture of his various strategies, and we propose to search for the best mixed strategy he can have.

Suppose he mixes his strategy $\boxed{1}$ and $\boxed{3}$ and plays strategy $\boxed{1}$ half of the time, and strategy $\boxed{3}$ half of the time. If it is assumed that I play strategy $\textcircled{1}$ x number of times, then my gain will be given by the formula

$$z = 1/2(7 - 5x) + 1/2(2 + 9x) = 4.5 + 2x \qquad (9)$$

In Figure 3 we give a graphical representation of this mixture of strategies $\boxed{1}$ and $\boxed{3}$. For instance, it can be seen that if I play my strategy $\textcircled{1}$ 80% of the time and he is mixing strategy $\boxed{1}$ and $\boxed{3}$ in equal proportion, then my gain will be \$6 (point A). We see, then, that if he plays his strategies $\boxed{1}$ and $\boxed{3}$ half and half, all he can be certain of is that he is going to keep me down to \$6.50, as I can realize this gain by playing the pure strategy $\textcircled{1}$.

Now we ask the question, what mixture of strategies $\boxed{1}$ and $\boxed{3}$ should he use in order to be certain to keep me down to the lowest gain? The graphical representation of each of these mixed strategies is given by a straight line going through point P in Figure 3. We see that the best mixture of his strategies $\boxed{1}$ and $\boxed{3}$ will be represented by the horizontal line in Figure 3. It is easy to compute that this horizontal line is obtained by the opponent playing his strategy $\boxed{1}$ 9/14 times, and playing his strategy $\boxed{2}$ by 5/14 number of times. In this case, we get for my gain the expression

$$z = 9/14(7 - 5x + 5/14(2 + 9x) = 73\ 1/4 = 5.2. \qquad (10)$$

This shows that my gain, (if he uses this particular mixed strategy) is \$5.20 independently of what strategy I play. (This means then, that the line in Figure 3 is really horizontal.) We can see then, that if he plays this mixture of his strategies $\boxed{1}$ and $\boxed{2}$, then he can be absolutely certain to keep me down to a gain of \$5.20. However, by inspecting Figure 2, we can see that he will be better off by mixing his strategies $\boxed{2}$ and $\boxed{3}$ instead of $\boxed{1}$ and $\boxed{3}$. More specifically, one can compute that by playing his strategy $\boxed{2}$, 9/11 times, and $\boxed{3}$, 2/11 times, he can keep me down to a gain of 49/11 dollars. This mixture of his strategies $\boxed{2}$ and $\boxed{3}$ would be represented by the horizontal line going through the point D in Figure 2.

In summary, we can see that if he plays this last strategy, then he can be certain to keep me down to 49/11 dollars. If he plays this strategy no matter what I do my gain is not going to be over the 49/11 dollars.

Let us remind ourselves now that previously we established that if I follow my best strategy I am certain to get the 49/11 dollars. Now, we find that if my opponent plays his best strategy he can be certain of keeping me down to this 49/11 dollars. Is this a coincidence that these two dollar figures are one and the same?

The celebrated theorem of the Theory of Games, the so-called min-max theorem, states that it makes no difference whether we consider the optimum strategy of either of the players, the expected payoff is the same. This theorem allows then the development of a Theory of Games and leads to a rational choice of strategies on the part of the players. Let us add, that a more precise statement of the theorem includes the specification that we are dealing with a "zero-sum two-person game".

## Concluding Remarks

We have presented here some of the basic concepts of the Theory of Games and, in particular, we have described what is meant by a "strategy". The Theory of Games is a field of science covering a large body of knowledge, and we have covered here only a very small portion of the theory. However, it can readily be seen that there is a similarity between "playing a game" and controlling a business firm under competitive economic conditions. In fact, the Theory of Games was originated to develop the economic theory of competition. During recent years, there has been a great deal of discussion of applying the Theory of Games to the solution of business problems. However, in most business problems, the knowledge of various alternatives is not well known, and, also, it is often difficult to predict the consequence of a decision. Also, conditions governing the behavior of business firms are usually much more complex than the conditions assumed by the mathematical theory. The Theory of Games is potentially an important tool for the business man, though actual applications are still difficult to find.

| | | His Policy 1 (1/3) | 2 | 3 | 4 | 5 (2/3) |
|---|---|---|---|---|---|---|
| My Policy | ① | 30 | 34 | 37 | 39 | 40 |
| 1/3 → | ② | <u>35</u> | 38 | 40 | 41 | <u>39</u> |
| 2/3 → | ③ | <u>39</u> | 41 | 42 | 40 | <u>37</u> |
| | ④ | 42 | 43 | 41 | 38 | 34 |
| | ⑤ | 44 | 42 | 39 | 35 | 30 |

TABLE OF MY PROFITS FOR EACH BARREL OF OIL.

FIGURE I. WHAT TO PRODUCE IN THE FACE OF COMPETITION.

AN APPLICATION OF THE THEORY OF GAMES. IF I CHOOSE POLICY ② AND MY COMPETITOR POLICY [4] MY PROFIT IS 41¢/BBL.

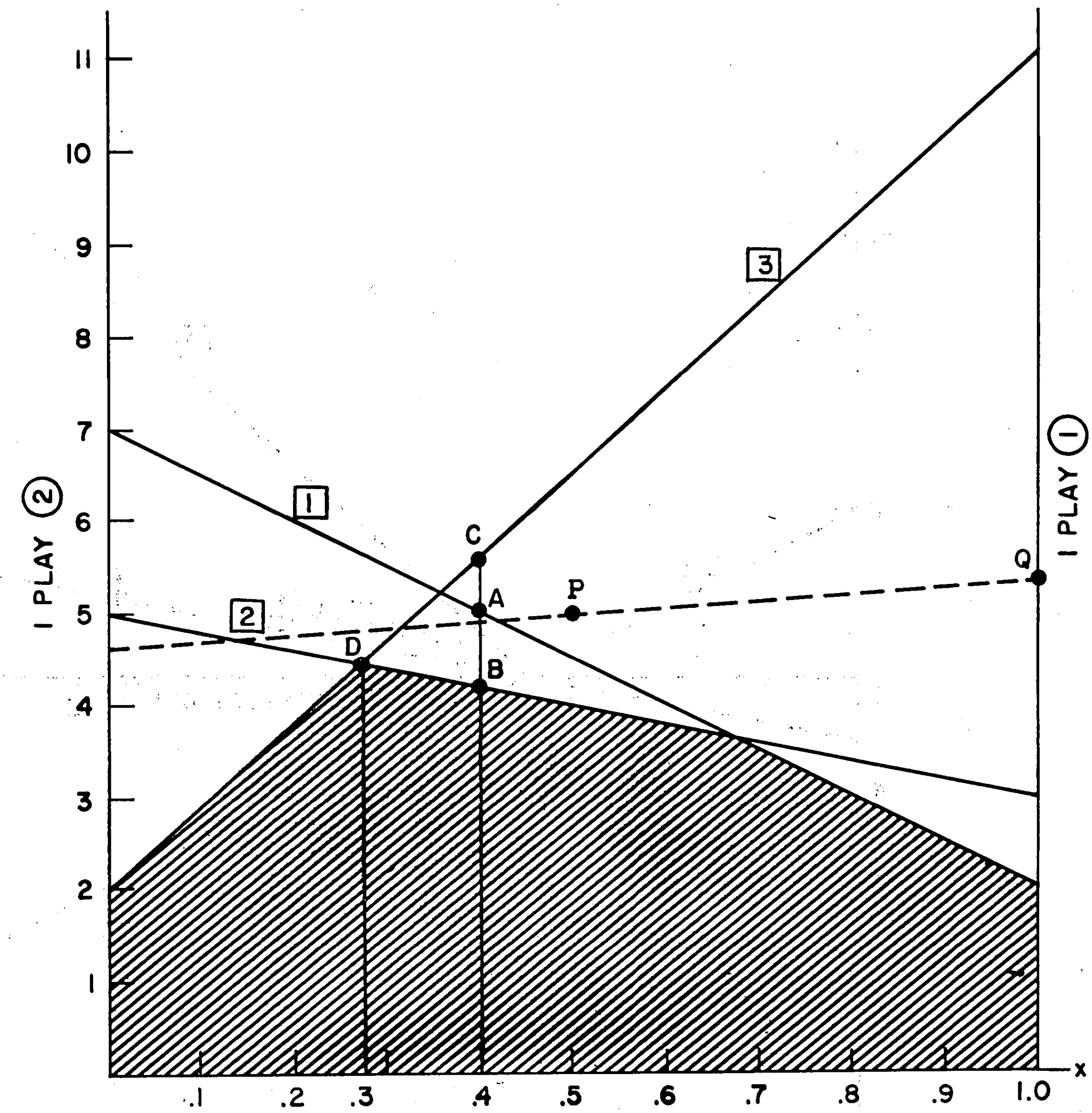



FIGURE 2. GRAPHICAL REPRESENTATION OF PAYOFF FOR ILLUSTRATIVE EXAMPLE. THE LETTER X DENOTES THE FRACTION OF TIMES I PLAY STRATEGY [1]

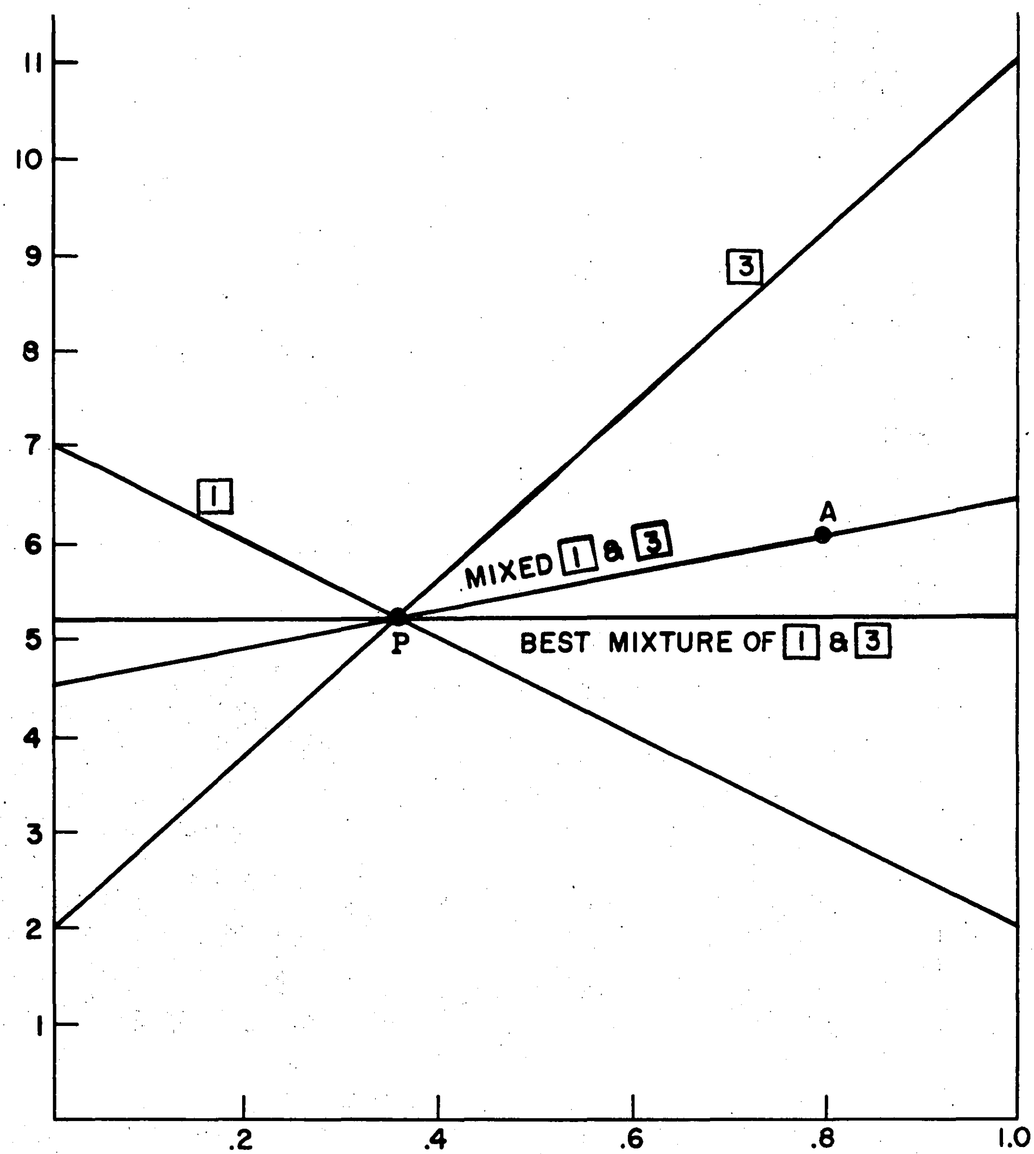



FIGURE 3. GRAPHICAL REPRESENTATION OF MIXED STRATEGIES 1 AND 3

102688959