Proceedings of the

EASTERN JOINT COMPUTER CONFERENCE

November 7-9, 1955 Boston, Mass.

THEME: COMPUTERS IN BUSINESS AND INDUSTRIAL SYSTEMS

SPONSORS:

THE ASSOCIATION FOR COMPUTING MACHINERY

THE AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
Committee on Computing Devices

THE INSTITUTE OF RADIO ENGINEERS
Professional Group on Electronic Computers

Printed in the United States of America

Price $3.00

PROCEEDINGS OF THE EASTERN JOINT COMPUTER CONFERENCE

PAPERS AND DISCUSSIONS PRESENTED AT THE JOINT ACM-AIEE-IRE COMPUTER CONFERENCE,

BOSTON, MASS. NOVEMBER 7-9, 1955

THEME: COMPUTERS IN BUSINESS AND INDUSTRIAL SYSTEMS

SPONSORS

The Association for Computing Machinery

The American Institute of Electrical Engineers Committee on Computing Devices

The Institute of Radio Engineers Professional Group on Electronic Computers

Published by

THE INSTITUTE OF RADIO ENGINEERS, INC.

1 East 79 Street, New York 21, N.Y.

ADDITIONAL COPIES

Additional copies may be purchased from the following sponsoring societies at $3.00 per copy. Checks should be made payable to any one of the following societies:

INSTITUTE OF RADIO ENGINEERS
1 East 79th Street, New York 21, N.Y.

ASSOCIATION FOR COMPUTING MACHINERY
2 East 63rd Street, New York 21, N.Y.

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
33 West 39th Street, New York 18, N.Y.

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 54-9071

JOINT ACM-AIEE-IRE COMPUTER COMMITTEE

PURPOSE: "The committee shall aid in the promotion of close cooperation and coordination in the activities of the sponsoring societies related to the field of computer engineering and allied arts and sciences. Its objects shall be scientific, literary, and educational in character and in harmony with the aims of its sponsors."*

Chairman

For the East
J. G. BRAINERD, University of Pennsylvania

For the West
O. WHITBY, Stanford Research Institute

Joint Secretary-Treasurer
W. D. LEWIS, Bell Telephone Laboratories

ACM Representatives

East
C. W. ADAMS, Westinghouse Electric Corporation
H. F. MITCHELL, JR., Sperry-Rand Inc.

West
J. BARNES, University of California, Los Angeles
D. H. LEHMER, University of California, Berkeley

Ex-Officio—A. S. HOUSHOLDER
Oak Ridge National Laboratory
President, ACM

Headquarters Representative—E. BROMBERG
New York University Institute of Mathematical Sciences
Secretary, ACM

AIEE Representatives

East
J. G. BRAINERD, University of Pennsylvania
D. C. ROSS, IBM Engineering Laboratories

West
R. R. BENNETT, Ramo-Wooldridge Corporation
P. L. MORTON, University of California, Berkeley

Ex-Officio—E. L. HARDER
Westinghouse Electric Corporation
Chairman, AIEE Committee on Computing Devices

Headquarters Representative—R. S. GARDINER
Assistant Secretary, AIEE

IRE Representatives

East
D. HAAGENS, Control Instrument Company, Inc.
J. R. WEINER, Sperry-Rand, Inc.

West
W. F. GUNNING, International Telemeter Company
O. WHITBY, Stanford Research Institute

Ex-Officio—J. H. FELKER
Bell Telephone Laboratories
Chairman, IRE Professional Group on Electronic Computers

Headquarters Representative—L. G. CUMMING
Technical Secretary, IRE.

* From the "Organization of the Joint Computer Committee" which may be found in 1954 Proceedings of the Eastern Joint Computer Conference.

TECHNICAL PROGRAM COMMITTEE

Chairman: IRVEN TRAVIS, Burroughs Corporation

J. C. McPHERSON, International Business Machines Corp.
H. T. ENGSTROM, Remington Rand
J. D. ELLIOT, Detroit Edison Company
E. F. COOLEY, Prudential Insurance Co. of America
C. W. ADAMS, Westinghouse Electric Corporation
D. L. STEVENS, Burroughs Corporation

EASTERN JOINT COMPUTER CONFERENCE COMMITTEE

1955

Chairman: J. G. BRAINERD, University of Pennsylvania

Secretary: H. F. MITCHELL, JR., Sperry Rand Corporation

LOCAL ARRANGEMENTS COMMITTEE

Chairman and Secretary:	F. M. VERZUH, Massachusetts Institute of Technology
Finance:	R. H. GREGORY, School of Industrial Management, M.I.T. W. J. OLIPHANT, Arthur Andersen and Company
Registration:	W. A. JOHNSON, International Business Machines Corp. CARL REYNOLDS, International Business Machines Corp.
Hotel Arrangements:	F. E. HEART, Lincoln Laboratory, M.I.T. P. R. BAGLEY, Lincoln Laboratory, M.I.T.
Inspection Trip:	R. C. KELNER, Laboratory for Electronics, Inc. A. E. OETTINGER, Harvard Computation Laboratory
Exhibits:	H. I. COHEN, Datamatic Corporation M. A. MEYER, Laboratory for Electronics, Inc.
Publicity:	J. C. HOSKEN, Farrington Manufacturing Company LAURENCE ROSENFELD, Melpar, Inc.
Printing:	D. T. ROSS, Servomechanisms Laboratory, M.I.T. ROBERT KRAMER, Servomechanisms Laboratory, M.I.T. ROBERT LEONARD, Servomechanisms Laboratory, M.I.T.

PUBLICATIONS COMMITTEE

All of International Business Machines Corp.

Chairman: N. P. EDWARDS

Committee

G. M. AMDAHL
C. T. BAKER
W. G. BOURICIUS
B. O. EVANS

R. A. GREGORY
R. E. MERWIN
G. E. PETRIE
D. C. ROSS

C. E. WALSTON

Secretary: MURIEL KOPP

TABLE OF CONTENTS

Keynote Address

J. G. BRAINERD†

AS CHAIRMAN of the Eastern Joint Computer Committee I add my welcome to that of Dr. Travis and express the hope that this meeting will be beneficial and stimulating to all of us. Before discussing the conference, let me give our thanks to the three committee chairmen, Dr. Travis of the Program Committee, Dr. Verzuh of the Local Arrangements Committee, and Mr. Edwards of the Publication Committee, to Mr. Porter who was chairman of the Local Arrangements Committee until his work took him out of this region, and to all others who have contributed time and effort to create this meeting.

As is customary, all in attendance will receive printed copies of the proceedings, and there it will be possible to acknowledge by name the many committee members who have worked to make the conference a success. Numerous others not formally associated with committees have helped, and to them also our thanks are due.

The Joint Computer Committee was set up early in 1951, and the first Joint Computer Conference was held in December of that year in Philadelphia. The Joint Computer Committee is a body sponsored by three technical societies to conduct conferences and carry out certain other work so as to avoid duplication of effort on the part of its sponsors. There are now held annual meetings both on the West Coast and in the East, and this is the fifth annual meeting of what is now known as the *Eastern* Joint Computer Conference.

Customarily, one major topic is selected, and invited papers are presented, building up the subject until a comprehensive coverage is obtained. I wish to emphasize that all the papers at this meeting are results of invitations to the authors, and have been selected to bring out principles, rather than specific applications, within the chosen topic of computers in business and industrial systems. In connection with the latter phrase, "industrial systems," the topic is not intended to cover the engineering-scientific-mathematical applications often discussed, but rather applications in the commercial and planning aspects of industry including such a topic as inventory control.

The purpose of this meeting is two-fold: first, to present for engineers and others in the field a background of fundamentals in order that they may better carry out their work of creating computers for business and industry, and, second, to present the same fundamentals to all others who may have an interest in using computers. It happens this year that our topic is such as to have an almost equal appeal to both groups. The program is so drawn up that it blends somewhat naturally, at the end, into a timely topic which should be of major interest to both creators and users, namely that of standardization of some aspects of the field.

In one sense it is fitting that I should be speaking now, for I was associated with the creation (at the institution with which I am connected) of the ENIAC, the first electronic general-purpose large-scale digital computer. The usual person's close association with the decimal system leads to an emphasis on integer multiples of ten, particularly with regard to anniversaries, and the present is almost exactly the tenth anniversary of the initial operation of the ENIAC. The ENIAC may soon be retired, but it has to its credit more computation than was performed for all the formal mathematical tables ever published in the world before its advent. This is no mean accomplishment, but mathematical tables and mathematical calculation are not our interest this week.

Many persons in the computer field will feel at least vaguely familiar with the time of start of the first big electronic machine; but few either in the computer field or interested in the application of computers to business will know that what was *probably* the first paper on the use of large electronic machines in business might have appeared in print almost at the same time as the inauguration of the ENIAC, had not reviewers considered it too "futuristic" for publication. The author was Professor Matz of Penn's Wharton School of Finance and Commerce, and it took him more than six months to convince the appropriate people that his articles, "Electronics in Accounting," based on his acquaintance with the ENIAC, was not unreasonable and should be printed in the Accounting Review (the journal of the American Accounting Association), where it appeared in 1946.

The first applications of the large digital electronic computers—we use the inadequate term computer for much of what is sometimes called data processing equipment or information transformer or another name, as well as for machines strictly "numerical" in intent—to repeat, the first applications were in the engineering-scientific-technical-military fields. In the ten years since their inception, computers of the type under discussion here have come to have important positions in control and the general area of automation, in the vast nebulous territory of Operations Research, and in many other classes of application, even including as you know a potential use in the translation of lan-

† Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

guages, but it now seems apparent that, at least as judged by number, the largest field of use will be in the business-commercial-financial area. In ten years the hesitancy which marked the reception of Professor Matz' paper has given way to optimism and acceptance, popular, semi-popular, and technical articles galore, and a feeling on the part of many that is most briefly expressed by the cliche, "the coming clerical revolution."

These ten years have seen discussion move from a few particular machines to talk of—and I emphasize those words "to talk of"—hundreds of the large computers, thousands of medium-sized ones, and an industry which may possibly be measured in the billion-dollar annual gross business class in the not far distant future. Estimates of personnel needs in the field which seemed extreme only a year or so ago appear much more reasonable now. Current optimism has its foundations, over and above such specific knowledge as actual orders on manufacturers' books, on one assumption and one contrast or challenge. The assumption is that prosperity and all that implies will continue in the United States. The contrast is between two basic facts, which will be labeled (a) and (b) for convenience.

(a) After allowing for cyclical changes and other factors, the number of production workers in manufacturing has not substantially changed in the last ten years.

(b) Against this is to be set a statement appearing in the magazine *Time*, which an economist advisor could not check immediately but which gives an idea of trend at least, that "the number of office workers has actually risen from 5,100,000 to 8,100,000 in the last ten years."

With increasing production, increasing population, and improving average standard of living, statement (a) indicates that production workers are increasing their per capita productivity—as is well recognized, the figure usually quoted being about two per cent per year, although recently there has been a tendency to predict a higher rate. Increased mechanization—or automation if you like—in which computers will play a more and more prominent part in the future, accounts for a large portion of the two per cent per year. Statement (b) indicates that the accompanying piling up of clerical detail, to which must be added the increased demand for services such as banking which require much clerical work, has not been even remotely overcome by whatever increase in per capita clerical productivity may have occurred. Here lies the opening for office automation (as well as for other methods of increasing clerical productivity), and here is the basis for *Time's* sister magazine, *Fortune*, to write of "The Coming Victory over Paper," with a heading inside the article, "Rescue by Electronics."

In the last ten years of our lives, which are the first ten years of the large-scale electronic digital machines, the field of computers has matured at a rapid rate. We have seen increased emphasis placed on peripheral equipment and on storage, and increased attention given to common and automatic programming, to cite only a few examples. There has been a relative transfer of research and development work from university laboratories to industrial research laboratories, except for numerous classes of special cases. There have been many technical developments and we know there are many more to come.

There has been a change in the center of gravity of application of the big and medium digital computers toward the business field. Our topic and program are a recognition of this. The program is so arranged that there will be ample time for discussion at most sessions, and it is our hope that we have avoided rush so that thoughtful attention can be given here, and in these three days, to the import of each session.

Computers as Tools for Management

J. S. COLEMAN†

IT IS A REAL pleasure to speak at this national meeting of electronic computer engineers. Each year your Joint Computer Conference has marked a new step forward. This year's conference seems to me to have special significance in that you are here to assess the capabilities of digital computers for handling business computations. I say "special significance" because I believe that in this area, where the computer becomes a tool for management decisions, there is a great potential for the growth of a new major industry.

In these circumstances, it would seem necessary to look to the fundamentals right now, early in the growth period. Most of the business applications of the computer, which we have planned up to this point, have led in one direction: greater speed in the processing of business data. To achieve this goal, we have aimed at more and more complete office automation. Probably that was a sound choice of a starting point. For business became interested in the elctronic computer when it was pointed out that the machine can effect great savings in the processing of data. Hence this application of the computer was the foot in the door needed to get this industry off to a sound start financially.

A number of effective applications of this kind are in sight, and the number of good applications will gradually increase. Although it seems over-optimistic to call office automation, to the degree that we currently foresee it, a "second industrial revolution," nevertheless, it appears certain that an evolutionary process is under way. It will clearly take a number of years to exploit the various possibilites over a wide variety of types and sizes of business.

But in spite of the encouraging start that has been made, is the long run significance of electronic computers for business to be found principally in their ability to process data, however tremendous a speed they may achieve? I think not. Adding machines and accounting machines introduced less spectacular but highly significant improvements in their day. And I dare say, as the mushroom cloud of this electronic computer explosion settles down, the power of high speed and routine data processing will not exceed the over-all power of the simple adding machine in cutting the costs of business. The inventors of the steam engine and the automotive engine will in the day of atomic power still hold their positions in the history books.

Considered merely as a data-processing tool, the computer is only one more cost-cutting device—and not yet completely proved, at that. Much more significant is the use of the computer in solving management problems of a high order. For, when it is employed as a tool for management decision-making, the computer achieves much more than its negative role of cost-cutting. It can then play an important part in decisions that will increase the revenue of the business and maximize profits. It is this latter and broader application of the computer that I should like to discuss, for here, I believe, lies the best hope for the future of the industry.

Perhaps I should pause here a moment to explain my use of the term "the electronic computer industry." At present you are faced, as electronic computer engineers, with a series of fascinating challenges. The challenge which you have clearly mastered is that of scientific and engineering computation. You are on the threshold of establishing some areas of effective business application. You are actively developing computers for control of military equipment and aircraft. You are groping toward use of computers in industrial process control.

With this variety of basically different and basically important applications, there is the prospect for emergence of electronic computers as a major industry. Although you are working in parent companies in the business machine field, the electronic and electrical machinery field, the aircraft field, and others, I do not think that you should any longer shy away from the term, *the electronic computer industry*. Of course, despite its rapid growth this is still an infant industry, not only in age, but in other important respects. Economically, as most of you no doubt know, it is largely a *dependent* of the parent companies which are creating it.

This year's Joint Computer Conference is intended to bring computer designers together with management users of computers for business applications. This tempts me to suggest: Why not bring together, next year or sooner, management and computer engineers to consider the economic aspects of our infant industry? From management's point of view I'm forced to ask: How long will we have the state of indecision as we go from vacuum tubes to germanium diodes, to transistors and bimags; is the end in sight in memory system progress from mercury delay lines, to electrostatic storage, to magnetic core memories before we operate as a true business? We are in the midst of a fabulous rate of technical development, but some degree of stability is necessary before electronic computers step out of the infant industry stage.

Alongside the need for establishing economic self-sufficiency, there is still the challenge of making computers truly useful on a widespread basis. At this meeting on use of computers for business applications, I think it is fair to say that although technical development of computers has gone far, the test of just how useful computers will be in business lies almost entirely ahead of us.

My purpose, then, is to examine with you the pros-

† President, Burroughs Corp., Detroit, Mich.

pects for a broader use of computers as tools for management information and decision-making, for it seems to me that it is this concept which constitutes a major challenge to the industry. We cannot completely separate this concept from that of routine data-processing. When one deals with the use of computers for accounting, the question of reports for management and the possibilities of better control by management are always present. But I submit that this is just scratching the surface.

Management already has more reports than it can handle. What it needs is help in exploring all the alternative solutions to its problems, so that it is better prepared to choose the best one. For this task, the computer is an important instrument. Let me illustrate what I mean. It is useful, of course, for the general sales manager to know that sales in the North Central Region are below quota for the month. And the computer is an excellent instrument to use in preparing the sales data report which gives him this information. But far more useful—in fact, of first importance—is the ability *to do something* with this knowledge, which is in itself quite sterile. Hence, a far more profitable use of the computer is in the preparation of alternative solutions to the sales problem. When the sales manager can see what each of several decisions involves, then he can choose the most effective one. Used in this way, the computer is not just part of a fire *alarm* system; it is part of a sound fire *prevention* program. This application cannot be regarded as a mere by-product of office automation. Rather, it is a major objective. What I am proposing is not something new, of course. It is, however, a redirection of emphasis.

Already encouraging signs exist that the importance of this kind of application is being recognized. There is the growing partnership between management, scientific research, and mathematics which is usually referred to as operations research. A few examples where operations research, management problems, and the use of computers have come to a point of common focus include the use of linear programming, the Monte Carlo method, and the game theory for optimum strategy. Much has been written on these subjects and I shall pursue them no further. In fact, if I did you might question me about the details and our staff experts are too far away for consultation.

Other mathematical methods which require extensive computations are finding important uses by management. For example, in the Burroughs Corporation, recruitment of personnel now relies quite heavily on the results of an extensive program of correlation of work performance with psychological test scores. Computations to determine the significant correlations were carried out on the electronic computer at our Research Center. Without such a computing facility, our program along these lines might never have been carried through.

As another example from the experiences in my own firm, we have made good progress toward setting marketing targets for our various products on the basis of national figures on the levels of employment and income in the major areas of business where we find our customers. The computations required to determine quotas for each of our branches have been a stumbling block. This year a small scale electronic computer solved the problem. Broadly speaking, these examples indicate that computers are becoming a valuable tool for some of the recent developments in scientific management.

There is a quite different approach, in which, although computers would serve as a data processing tool, the key objective would be improvement in management controls rather than the automation of clerical operations. The objective of some firms which are interested in electronic computers is to use the computer to permit complete integration of their market forecasting and actual sales activities, with production and inventory control. This is a truly challenging prospect in those types of business in which a high proportion of the management functions must go into line activities involving production and inventory control on the basis of sales prospects and orders in hand.

An electronic computer for handling management control functions would provide a facility for storing records of sales forecasts, sales orders, production capabilities, and inventories, and converting this information to production and shipping orders. The computing machine would be able to apply the best available industrial engineering techniques such as well-known production log size formulas and newly developing production and inventory planning procedures to determine the best production plans. Success of such an approach will require not only the electronic computer and extensive research on the computing program, but also an effective system of rapid communication between sales order points, production facilities, and the computer. For this reason those companies which are striving for this goal of computer use have shown especially great interest in the development of *integrated data processing* in which all types of business records provide for coded output, such as punched tape, to facilitate long distance communication and input to computers.

A concept of this sort cannot, of course, be envisaged in all types of business, any more than the concept of office automation with high speed computers applies to the handling of all business records. The first prospect for use of computers along these lines would be in large scale manufacturing in which alternative product models, or a line of similar products, share the use of manpower, raw materials, and capital equipment. The automotive industry, producers of small machine parts such as screws, bearings, and gears, and the radio and television industry, are examples. Large scale retail trade offers similar possibilities, without the element of production control, but with tremendous problems of inventory control.

These applications have to do with *system control* rather than automatic data processing. And how much more it means to business than mere cost-cutting, important as that may be. Let's say that a company uses an electronic computer just for the sake of saving money. We'll suppose that it uses the increased speed effectively and saves $200,000 a year. That is a sum none of us will look down our nose at. But let us say that this same company puts its computer to work at determining the number of salesmen necessary to cover a territory effectively. Or suppose that it uses the computer to arrive at a decision that a regional sales set-up is better than a national set-up with all control from the home office. It is not difficult to see that these latter two problems and their solutions are vastly more significant to that company than its economy measures. For these latter two affect the growth of the company, its revenues, its whole competitive position. Used in this way, the computer can help maximize profits and reduce expenses in a very basic manner. It could become part of the very lifeblood of the company, for lack of which a company could fail to survive.

A further indication of this fact is the manner in which operations research was born. In the Battle of Britain the military turned to science for help. Cost-cutting was of very minor significance at that time when the country was fighting for survival. This science was employed to arrive at the optimum use and distribution of men and material. And it was the outstanding success of this critical effort that led to the interest which business now has in operations research and its tool, the electronic computer.

It is interesting to note, incidentally, that the use to which the equipment is to be put must determine the type of equipment selected. Routine data processing and improved decision-making may be conflicting objectives if the same equipment is chosen for both purposes. This fact again indicates the importance of setting our aim on the proper target.

I have discussed two distinct concepts of an important role for electronic computers in management analysis and control: first, as an analytic tool based on operations research and other aspects of scientific management; second, as a system control tool bearing on production, inventory, and the related phases of industrial operations. Both approaches include large elements of speculation, but are based on significant trends and offer realistic challenges. The analytic concept suggests that electronic computers are linked with general developments toward scientific management, and that the time may come when at least the larger companies find it desirable to have a computer available for analytical computations on management problems in the same way as they are currently finding it desirable to acquire computers for engineering computations. The system control concept implies application only to certain types of business, but in those areas the computer may become an essential tool rather than merely a desirable one.

Before proceeding further with these ideas, it will be well to introduce a third line of evidence that computers are on the way to becoming tools of management—the various activities along these lines by government agencies, primarily by agencies of the Department of defense. The problems of military logistics are by-and-large the equivalent to many of the problems of business management. Since the end of World War II each of the military services has been carrying on extensive programs of logistics research, in which electronic computers have served as a major tool.

The Navy has sponsored an active "Logistics Research Project" by contract with George Washington University in Washington, D. C. The Air Force has a Directorate of "Management Analysis Services" in the Pentagon for intimate contact with Air Force planning problems, and sponsors longer range logistic planning research by contract with the Rand Corporation, in California. The Army carries on logistics research through its operations research program under contract with Johns Hopkins University. In each of these logistics research programs an electronic digital computer is about the only major piece of research equipment. The Air Force has placed a digital computer in the very bowels of the Pentagon for planning calculations.

The logistics research programs are parallel to the prospective use of computers as a tool for general business planning and analysis based on developments in scientific management. At the same time the military services have been exploring the possibilities for use of computers in the day by day control over the operation of their supply systems, an approach which parallels the possibilities for business system controls over production and inventory. The Air Force has begun a careful program of trial and analysis of use of computers for control over movement of Air Force supplies, and the Navy and Army have similar programs.

The leadership being taken by the military departments in these developments is parallel to an earlier period, when ballistic calculations and other technical military problems paved the way for the initial development of electronic computers and paved the way to their widespread use in science and engineering. In the area of management computations, military research programs have again assumed leadership. The reason is clear: human lives may depend on it. We are reaching the point where we want to know whether corporate lives may also depend on it. There is little doubt that industry, like the defense agencies, will take steps to find out. At this stage it is the research spirit toward logistics by the defense agencies which has done most to pave the way to use of computers in management analysis and control. If industry develops similar research programs on a fairly widespread basis, we can hope for rapid progress.

If the problems to be solved were a relatively straightforward matter of further development, we could speak with greater confidence about the rate at which we can progress. At this stage it is most important that we recognize the underlying problems and foster the research to solve these problems. In the first place the complexity of business management makes careful research necessary in order to determine the proper scope of the targets for application of scientific methods and corresponding use of computers.

In management system controls the problems are more complex than in the more purely technical system problems such as industrial process control. Management problems will require more extensive research to isolate the key factors, the interrelations, and techniques for obtaining good solutions from an environment of innumerable human factors and from broader management considerations which cannot be expected to fall within the scope of a scientific management approach.

An equally fundamental, but related issue lies in the uncertainties of making predictions which surround the problems of management. The very word *management* implies decisions which involve prediction of the outcome of alternative courses of action. I have recently seen the difference between human decision making and the choices made within mechanisms such as electronic computers nicely clarified by the statement that successful human decisions can be and generally are made without complete information, in fact, often with very scanty information. A machine, on the other hand, can make choices only on the basis of the information which it contains—it can only deduce consequences from whatever information has been fed into it.

This way of putting it seems to imply that the machine can only be a data processing tool. But, I am by no means retreating to that point of view. The machine can only be a tool, but a distinction should be made between a tool for clerical work, and a tool for management. The underlying question is whether a scientific rather than a rule-of-thumb approach can be found for some of the areas of management prediction. The approach cannot be a science involving certainties, such as is generally the case in the physical sciences. Instead, it must be a systematic method of determining possible outcomes, the relative probabilites of these outcomes, and methods, for basing decisions on the probabilities.

Research on business prediction and on the evaluation of uncertainties has been going on in recent years. The results have provided some useful guideposts for management, but do not as yet provide a general approach to any one well-defined area of management. We can do a fairly scientific job of market sampling on particular products; we can assume, or hope, that forecasts of the levels of general economic activity made by government and industry economists are based on increasingly scientific methods; we may find linear programming or other results of operations research applicable on some production and inventory problems, but these are at present isolated techniques, which, on the whole, deal with factors in management situations rather than the management situations as a whole.

Since any prospective science of business management will have to be a science for dealing with nothing but uncertainties, we have the further problem of determining whether any proposed scientific approach is actually superior to management decision on the basis of experience and intuition. We will perhaps have to go on a conviction rather than proof that the scientific approach is the desirable one, just as we go on conviction that the weather man is more often right than wrong.

That may be enough for us, but it may not be enough for the business man who is the prospective customer. He must be taught to have confidence in operations research and the part which the computer plays in it. He is from Missouri and he must be shown that the new science is superior to his intuition. May I suggest that in dealing with him, you be neither discouraged nor sarcastic. It is essential that management learn both to understand and trust your findings. The best way to bring this about would seem to be to explain what you are doing simply and in language he can understand. Of course, he need not understand the intimate details of either electronics or operations research any more than he needs to understand how the tax expert or controller arrived at his figures. But you must remember to go slowly with him, and to show that you are interested in his problems. If some areas of business decision can thus be brought into a scientific framework, then the prospects will be good for electronic computers to become important management tools.

These comments on the types of research which are going on and which lie ahead with regard to use of electronic computers as "management computers," have not brought us to any questions of computer system design. To some degree this is an expression of the widespread feeling that our ability to design electronic computers has run far ahead of our knowledge of how to use them.

Generally speaking operations research approaches a phase of management or a business system as though it could be analyzed by the same general methods as are used on engineering analysis of technical equipment. This suggests that the use of computers as management tools is likely to have a good deal of resemblance to their use as tools of engineering. If scientific methods for analysis of management problems make substantial headway, then computers will be important for solution of those management problems, just as they are for solution of engineering problems. If operating controls over certain phases of business by means of computers are based on technical analysis of business systems, that will be quite similar to engineering analysis of complex machinery in order to use computers in controlling technical processes.

In the important developments of scientific management undoubtedly computers have been an important stimulus. Technical methods may be as important in problems of management as they are in problems of engineering, but the problems in management are more complex. The possibilities of actually working out solutions became much more promising when electronic computers became available. Thus, electronic computers have helped pave the way to a new spirit of research on management, just as they have brought us into a period of research on methods of business data handling.

In my opinion the concept of "computers for management" is a challenge well worth considering. It is a large undertaking in which the computer industry can proceed only in partnership with management analysts and with management itself. If the computer industry provides leadership, I believe it will be taking a long step toward opening new markets for its products.

Computers in Basic Business Applications

F. J. PORTER, JR.†

IN APPEARING before you to tell you of the trials and tribulations of the businessman who tries to use the equipment you are designing and building for basic business applications, I presume I should qualify myself as an electronics expert. A friend of mine, struggling to make one of the big machines work on a business problem, defined an electronics expert as "a man who hasn't yet received his computer." Under this definition I qualify fully because my company, the Consolidated Edison Company of New York, has not yet received the two Univacs and the one 705 it has on order. Next April after they are delivered I presume we shall appreciate to the utmost the point of his definition. Today, ignorance is bliss.

Since we are not yet in operation at Con Edison, my personal experience upon which these remarks are based only goes through the early study, decision, organization for programming and initial programming stages. The problems of operation are still ahead of us. Nevertheless, I believe I can show you enough about these initial problems to cause you to realize that problems do not stop with the design and construction of a computer but are just as severe and just as baffling in some cases on the application side of the picture as they are on yours.

Then I shall try to outline a few areas you might want to study from the business angle so that the equipment you produce may be of utmost value to the user. I believe that one of the greatest needs of the computer industry today is a closer liaison with the user, for you can only sell that which is of some benefit to him and to produce anything else is an economic waste.

The Businessman's Problems

The electronic digital computer burst upon the business world some five or six years ago in a rash of articles in the magazines directed to business management. "Giant Brains" were the beginning of "Automation." "The Automatic Office" was certain to make people obsolete. The day of the millennium had come! These startling and breathtaking concepts, akin very much to the dreams of the future which appeared when the atom was first split, caused quite a stir and presented us with the first problem that every businessman faces when he begins to look at this field. He must find out what the shouting is all about.

Study and Training

The first step in this is education. He knows right at the start that someone must decide whether or not to proceed into office automation. Since the decision must be made with judgment, foresight, and care, the problem cannot be left entirely to junior members to study. Someone in a fairly high place in the organization and in whom the top management has confidence must take the time to acquaint himself with the problem and the equipment. In many companies this has taken the form of a committee, either of middle management personnel reporting to a top executive or of top executives guiding specially assigned middle management personnel. This latter is the approach used in Con Edison. Two of us spent a considerable part of our time for several years studying the field and reporting periodically to a committee of Vice-Presidents and associated executives.

This process of self education for the businessman starts with the reading and studying of everything he can find in print. He will then talk with the manufacturers' representatives at great length and perhaps visit actual installations. At this point he will probably feel that he still needs to know more about the detail if he is to be able to understand and evaluate not only the manufacturers' claims but also the problems his own staff will bring to him, for he soon will see the need for full time personnel in this field. He may want to enroll in the manufacturers' programming courses, especially the shorter ones, and thus gain at least a working knowledge of equipment and programming techniques. This knowledge will enable him to guide his staff with some degree of intelligence and will help him to listen more critically to the basic sales pitches of "fixed vs variable word lengths," "internal vs external checks," "plastic vs metallic tape," and so forth.

It must be obvious to you that this is a time-consuming task and that it is no small or insignificant problem the businessman faces when he approaches this field. Yet the pot of gold which seems to be at the foot of the computer rainbow appears to be so large that it cannot be neglected. He must take the time and make the effort to learn all this even though he may sometimes wonder privately whether it will be as big as he thinks when he reaches it.

Businessmen or Scientists

Early in the approach to the automation problem, the businessman is confronted with a jumble of technical terms more or less meaningless to him. "Bits," "binary adder," "flip flop," "minimum latency," and a host of others dance before his eyes and cause him to wonder whether this whole subject isn't just too technical for him to master and wouldn't he be wiser to retire gracefully and leave it all to the engineers. Most businessmen are not engineers and those of us who are, are of a vintage when Mark I was not even a gleam in

† Commercial Manager, Consolidated Edison Co. of New York, Inc., New York, N. Y.

Doctor Aiken's eye and the best vacuum tube we had was a little thing called an "Audiotron" which had no base but was suspended from binding posts on the front of a panel by the leads to its elements.

After a while this mood of dark despair passes as further study and thought shows that engineers are not needed to apply this equipment to existing business problems. We do not need to know how the machine thinks or the language in which it thinks so long as it hears and speaks in a language intelligible to us. And since both the input and the output of most computer systems are in familiar alphanumeric characters, the businessman can communicate with the equipment relatively easily.

Furthermore, a much more fundamental reason soon appears. Digital computers of course were first used for scientific purposes and from this grew the legend that only scientists could understand them. The truth of the matter is that scientists were needed to understand the problems to be solved and not the internal operations of the machine, which really do not need to be thoroughly understood to use them. Reasoning from this, the businessman suspects that since he has business problems to solve, he should use businessmen who knew the problem thoroughly, to solve them. Now I do not say that engineers or mathematicians might not do the job better, provided they knew as much about the business problem as did the businessman. But I am satisfied from our experience that it is entirely possible to train businessmen who have had extensive experience with specific business problems how to use a computer. From what I know of business problems I suspect that this will be easier than to try to teach computer experts the complexities of business.

Is There An Application?

About the time he has reached this stage of his development, the businessman begins to face up to his next problem. Does he have an application which can economically use a digital computer? If he does, then he should build up a staff as quickly as possible because they too must go through a time-consuming period of training before they can tackle the procedural and programming task. If not, he should reduce the amount of time he is spending in the field to just enough to keep abreast of new developments which may change the picture. How can he find out quickly without making too elaborate a study?

Here is where his judgment, experience, and careful thinking will pay off. He must pursue several different lines of investigation and then reach a decision. A quick survey of the areas that might be affected within his own company will tell him whether or not he now incurs sufficient operating expenses to support the cost of proposed equipment out of savings in these expenses. The same knowledge of his operations will tell him whether he is presently highly mechanized or whether there still remains much manual work which could be done on an electronic data processing system. Talking to manufacturers' representatives can be of help, especially if they are familiar with his business, but he must keep a little salt handy just in case the salesman's enthusiasm outstrips his good judgment. He can inquire as to what others in his industry have done in the same area and what other industries have done in the same business function. Having accomplished all this, he should then dream a little about what untapped potential there may be in the way of faster and better reporting, centralization of operations and radical changes in method. Now he is ready to decide whether or not he has an application for a digital computer.

If the answer is yes, the next problem confronting him is to make a careful economic evaluation of the application and the possible types of equipment to handle it. This includes the preparation of procedural charts to show how the major flow of work will take place; a study of the personnel and equipment requirements; and a consideration of the subsidiary problems encountered. This is obviously a careful, lengthy task and suggests that now is the time to assemble his staff if it has not already been done.

Selection of Staff

The fundamental problem of staff selection is of course, who? Should he go out into the highways and byways and hire young engineers and mathematicians (if he can get them) or should he be content with older, more experienced methods men, who have grown up in the business but have a minimum of formal training. The first approach gives him a bright, enthusiastic, eager group fresh from study but lacking in business know-how, while the second provides plenty of experience but perhaps a lesser enthusiasm for change and a less recent exposure to organized study. Again he reverts to his previous fundamental decision that knowing his business is all-important even at the expense of some enthusiasm, and draws the bulk of his staff from experienced methods and supervisory personnel. However, to prevent the group from being too conservative he leavens it with a sprinkling of young men, drawn if possible from his own organization. We have found at Con Edison that teaming young and old makes ideal programming groups, for the experience of the older man and his knowledge of the business act as a brake on the impetuosity of the younger man who in turn brings an enthusiasm for new and untried methods which stimulates the older man. He needs, however, to place in charge of this staff a man of broad experience in the particular business operation being studied and one who not only has a flexible and open mind but is a good organizer as well.

Buy Now or Wait?

As the staff proceeds through training to the economic study, the businessman begins to think about his next problem. How much weight must he give to technological obsolescence? Or to put the problem in his terms, "Should he jump in and get a computer now or wait until there are further developments in the equipment field?" Obviously, the field is going to change and obviously better equipment will be produced. But by waiting he prolongs the day when savings will start to accrue. If he plunges now he will probably have to change again in five years and incur additional changeover costs. If he waits, the industry may always dangle before his eyes the prospect of better equipment to come. Somewhere along the line he must decide to go ahead, foregoing the possibilities of the future for the realities of the present.

Here again his judgment probably will give him the best guide. Many firms who see immediate profit in using computers have gone ahead now, fully realizing that sometime in the future it will all have to be done over. Many of us feel that the experience we gain out of proceeding now rather than waiting is worthwhile even if we only break even now, because we will be in a better position to evaluate the equipment that is to come. Furthermore, by going into the marketplace now, we are encouraging the manufacturers to proceed with future developments. We realize they have to sell equipment if they are to have the resources to improve. Some companies feel differently, however, and each businessman must decide the technological obsolescence question for himself.

Re-Think?

Now that his staff is selected, trained, and hard at work on an economic study, the businessman feels he can relax for a time and turn his attention to some of the other problems of his job. Unfortunately he is soon disillusioned. The staff begins to bring him basic problems beyond their authority to settle. One of these is the re-thinking problem. It is a pet belief of many working in the computer field that this equipment is so new and so different, with so many exciting possibilities, that totally new concepts of data handling are necessary to use the equipment effectively. These people contend that we should re-think our entire data handling process to secure the ultimate advantage of computer techniques immediately.

There is another school of thought which seems to me to be the prevailing one of the moment. Re-thinking is a time consuming process. It is creative work, and creative work is usually slow. Furthermore, it not only requires the development of the new things we want to do, but means that we must secure acceptance of completely new concepts by our organizations. This may be a formidable stumbling block. Where there appears to be an immediate saving possible by applying computers to present procedures with only such changes as are inherent in the equipment or are relatively easy of accomplishment, many businessmen will prefer to take the evolutionary rather than the revolutionary approach. After they and their organizations have experience with the new equipment they feel they will be in a better position to re-think the data processing problem. This is the approach we took at Con Edison. It is our intention, once we are in operation, to review our accounting system design carefully to see whether a more thorough revision is then justified.

Extent of Reorganization

No sooner does he decide this question than the staff asks the businessman how far they should go in planning on reorganization of the functional structure of the company. They point out that the big computers thrive on volume and if everything can be brought to one place under one direction, the computer can operate most efficiently and with the greatest money saving. Now the businessman really has a problem on his hands. What the staff says is true, yet to reorganize may cut across many long established lines of executive authority and affect many organizational sacred cows. The businessman can find no ready-made answer to this problem for it is different for every company. Each must solve it as best he can in the light of the conditions he is working under. But you may be sure he will step lightly and move carefully before he gives the staff an answer. How did we solve it at Con Edison? We were already highly centralized and we have tremendous volume, so we did not have to face this one.

Invisible Records

From his own private study of electronic data processing, the businessman knows that his basic business records are going to be in the form of invisible magnetic dots on long pieces of magnetic tape. This worries him right from the start of his studies. How will he refer to his records and how fast can he secure necessary information? He knows he can't keep the customer—or the president of the company—waiting too long for answers to questions. He knows that his internal auditors will want to see records of some kind and that government auditors of various types will need to be satisfied. And yet he also knows that there is no electronic device on the market which will give him access to these data fast enough and cheap enough to solve his problem. To top it all he is still just a little bit uncertain himself about this idea of having everything invisible. Suppose something goes wrong—he won't even know the data has been destroyed. Or suppose he has to take his records to court, will the judge accept a roll of magnetic tape?

All of these doubts and needs drive him early to make the basic decision that there must be frequent printouts of information of sufficient completeness to enable clerks to answer questions, display records and audit books. The magnitude of this printing job bothers him a little when he begins to add up the lines to be prepared each day. For example, in Con Edison for payroll and revenue accounting alone it will amount to about 650,000 120-character lines per day. But he puts it down as one of the costs of the job and starts to press the manufacturers for higher speed printers to accomplish it as cheaply and quickly as possible.

Preparation and Programming Costs

As the staff delves deeper into their study, they begin to bring up more problems. The heat load in this equipment is tremendous and heavy air conditioning expenditures are going to be needed. Some air conditioning must be run continuously to maintain constant humidity. Power usages are large and in some cases voltages of nonstandard types and of unusual accuracies are necessary. The architects must be consulted to make sure that the proposed location of the equipment is strong enough to bear the weight. He finds he is definitely out of the electric typewriter and keypunch area and into the heavy equipment field. The costs to prepare the site loom large and cause some raised eyebrows when mentioned in high places.

The businessman has long realized he must provide for considerable programming cost. His early study in the manufacturers' courses showed him that the language of the computer is so different from anything they had known before that he must be prepared to have his staff spend much time in learning it and learning how to apply it in an efficient manner. So while he is prepared for large numbers of man years in this area, he is somewhat startled when the staff points out that analyzing, charting, and coding—the essentials of programming—are not the whole problem. Much fundamental work must be done in the design of the accounting system itself before programming can be started. All procedures must be reviewed to be certain that every minor detail is either provided for or omitted for cause. All forms must be reviewed and altered where necessary. New ones must be designed. New and original control procedures must be developed so that he knows that the process is operating properly and is not failing to record sales somewhere along the line or is paying employees too much or too often. A plan for converting his present operations without too much disruption of normal business must be developed and the people to man it must be secured. Finally, new procedures must be written and the employees who are to operate the new system trained in their duties. Our businessman sighs, picks up his papers, and walking into his superior's office, says "Well, Joe, it looks nearer to forty than twenty man years to program that computer for revenue accounting."

As he sees what his staff must go through to prepare for computer operation, another horrible thought goes through his mind. What will be the problem of keeping these procedures up to date and making changes in them? Business operations are never static. Laws change and new methods must be developed to comply with them. Just as an example, ask your Payroll Department management what they had to do to comply with Internal Revenue Department rulings when sick pay was declared exempt from income tax. Competitors come up with new ideas, and methods must be devised to meet them. Top executives want new or additional figures or figures on a little different basis. These things all must be accomplished in relatively short periods of time.

The businessman talks to his staff about this and secures little consolation. It looks like a problem to them too but hopefully they tell him that perhaps after they are more experienced they can make alterations in programming faster and more accurately. Certainly he will need a permanent force of trained programmers working constantly on his problems. Perhaps he will just have to tell that Vice-President that he can't have changed figures as rapidly as he would like. The businessman sighs again and mutters to himself that he will cross that bridge when he gets to it.

Machine Availability

In talking details with the manufacturers' representatives, the businessman finds he must change some of his concepts of machine availability. He is used to the virtually 100 per cent availability of typewriters, adding machines, billing and even punch card machines, with short maintenance periods once every three or four months. Suddenly he finds that anything better than 80 per cent usable time is looked on with pride by the computer people and he must give up his machine on almost every shift for appreciable amounts of preventive maintenance. He can foresee difficult scheduling problems if he is to accomplish his results on time with the interruptions which begin to be apparent. He now realizes that he must add considerably to the time estimates his staff is preparing, to allow for this problem and probably will have to count on overtime or partial additional shift operation as a more or less regular thing.

Problems continue to come to him from all sides. Having a detailed economic study from his staff showing an advantage in using computers, a decision must be made as to which manufacturer to patronize. This is not always obvious from the study. Many sales pressures will be brought to bear and many of the conflicting claims of the manufacturers must be evaluated before he can make up his mind.

When a decision is reached, the staff must be expanded to handle the over-all job. The same problem of "who" presents itself but on a larger scale. It is really beginning to hurt the organization to take our more experienced supervision, and internal organizational problems must be met. If the work force affected by computers is organized, the Union must be informed of the decision and suitable negotiations must be undertaken in line with normal Company-Union policies in such matters. Sometimes these are very time-consuming.

Areas Needing Study By Designers and Researchers

I could continue at some length but I believe I have said enough to convince you that the businessman user has problems just as baffling and just as difficult to him as many of your problems are to you. Now I would like to strike into the realm of opinion and point out those areas where I believe the designer and researcher might work to make this new equipment more usable. Naturally the opinions expressed herein are my own, based on my experience alone, and do not necessarily reflect the official views of my company. I believe, however, that my associates in Con Edison are substantially in agreement with them.

Programming Costs

The first one is the programming area. Many of us are appalled at the size, and of course, the cost of the programming problem. When we see the necessity of thinking in terms of 40, 50, and even 70 to 100 man years for programming the initial operation and the establishment of a fairly large force of permanent programmers we become disturbed. These represent investments which only the largest companies can possibly afford. Unfortunately, we see nothing today which indicates that programming has any relation to the size of the user. It appears to be practically as expensive to program a 5,000-man as a 25,000-man payroll, providing the complexities are the same. The same detail must be worked out, the same analysis done, the same charting and coding. Only the volume or time on the computer is different. How can the smaller company afford an electronic computer?

Now I know, of course, that work is being done to help this problem. Automatic coding may supply some of the answer but it only attacks a small part of the programming operation. It still seems to be well ahead of us on any kind of universal business basis although some individual firms appear to be making headway on compilers for their own particular problems. Perhaps it can never be general but must be tailor-made for each application. If so, this will greatly limit the usefulness of automatic coding.

Another point may help the programming problem. Much of our time is being spent now in the fundamental design of the business procedure being programmed. The programmers cannot analyze, chart, and code until they have something to work on. Much pioneering must of necessity be done in developing basic procedures to accomplish business operations such as preparation of payrolls, billing, stock control, purchasing, etc. Once a number of these installations are developed and in operation, parts may be adapted by others without all of the basic soul-searching now necessary. Perhaps through the good offices of the manufacturers there may be a better interchange of this basic system design data. But I fear this hope presupposes a greater degree of uniformity between companies in accomplishing basic business operations than now exists or may practically be expected in the immediate future.

There are possibilities in special purpose equipment with at least a few built-in program steps but this approach, too, has disadvantages which I shall cover later. Perhaps a study of the instructions you have built into the equipment would be of assistance. Are they all necessary? Are there enough? Which ones are seldom used? Could several be eliminated by some new instruction or could programming be simplified by a different type? Are the instructions needed in one industry the same as in another? Is it possible to build in special instructions or omit others without prohibitive costs? I have heard of nothing that is being done along these lines at present.

Physical Size

Let us leave programming now and consider the physical size problem. The machines you are making are just too big physically. They generate too much heat; they use too much power; they're too heavy. Peripheral equipment components have the same difficulties. The need for four boxes taking up several hundred square feet in order to print on paper seems hard to understand. Why cannot at least the power supply units be included with the control units and the tape handling mechanism be built right into the printer? Probably transistors will help the power-heat-weight problem, but where are they? Four years ago they were going to replace the vacuum tube but we are still using tubes. Can printed circuits help this problem? What is being done to reduce the number and size of components?

Perhaps part of the trouble again is the general purpose concept as against the special purpose concept. If equipment could be produced in small units which could be assembled by plugging or in some similar manner, we could have only what we need for a particular application whereas now we have to take everything or nothing. I seem to have a dim recollection, however, of a line of radio equipment produced many years ago, I think by De Forest, which had the same idea but which never became very popular. Whatever the solution, this is an area needing consideration.

Lack of Standardization

The next area which seems to me to need study is the lack of interchangeability of equipment or to say it more correctly, the lack of standardization within the computer field. Why shouldn't the magnetic tapes of one manufacturer run on the tape mechanisms of the others? Why shouldn't the various computer codes be alike so that a tape from a Univac could operate an IBM high speed printer, or 705 tape punch cards on a Sperry Rand tape to card machine? And why shouldn't tapes all be alike instead of some being made of plastic, some of metal, and some of paper? Isn't one of them just out-and-out better than the other?

I suspect that you may answer me by saying that these are more business than design problems because they enter the field of sales policies, patents, economics, and so on. To some degree this is true, of course, but you also bear some of the responsibility by insisting on the last fine measure of perfection in technical detail rather than compromising some of the perfection of your design to accomplish uniformity. Perhaps a slightly less efficient computer might be more useful to the buyer if he were able to use it with a competitor's piece of peripheral equipment which is better than yours or which you have not even produced yet. I believe it is a serious question whether all computer manufacturers can produce complete lines of equipment so that all business operations can be performed without combining the equipment of several makers. I even doubt the over-all industry-wide economy of trying. If there were a greater degree of standardization among some of these essentials, it would be possible for each manufacturer to restrict himself to those machines he happens to be best fitted to produce without having to spread himself ruinously over the whole field. If some thing like this does not develop, the businessman is almost forced to restrict his purchases to the products of a few large manufacturers who have the resources to build all necessary components. I suggest that these thoughts be kept in mind when you are tempted to argue for technical perfection at the expense of interchangeability.

Random Access

Earlier in my talk I mentioned the shadow of doubt which exists in the businessman's mind about invisible records and the need for random access to them on a high speed basis. This is still one of the weakest areas of the new electronic systems of data processing and much more work needs to be done in it. The random access devices now on the market are a step in the right direction but they are quite inadequate. Several million alphanumeric digits per machine are scarcely sufficient. We need to think in terms of a hundred million digits or more per machine. Con Edison's customer history file will contain almost a billion and one-half digits for some 2,700,000 accounts and you can imagine what an insurance company with five or ten or even twenty million policies would need for premium billing.

Here I see no technical answer even on the horizon and I must leave this problem to you. I do feel however, that we in business have not given you quite a fair mark to shoot at because we are not settled in our own minds just how many digits of information we need to store nor how fast we need access to them. The needs quoted above for Con Edison are based on five hundred digits per customer yet already one utility is talking about storing as few as two hundred. Perhaps experience will show us how we can get along with even fewer. Access time is another difficult problem to answer and no one of us is willing to commit ourselves on it. After we have had more experience in both of these fields we can, no doubt, give you more realistic specifications to guide you.

Sorting Techniques

Most business applications have one or more groups of input documents which come to the computer system in random order. For example, pay checks and dividend checks are returned by the bank mixed up, meter readings and cash payments are in no special order, and purchase orders do not follow the stock sequence. All of these must be placed in the sequence in which the data are carried on tape if they are to be posted speedily and economically. Usually the control is a number of some sort such as an employee number, account number, or stock number. The input data must be sorted either before putting on tape or after. Both methods present problems. Sorting before presupposes that the document has some automatic sorting medium in it such as punched holes or readable dots and is more or less of standard size. Sorting after transfer to tape is slow and wastes computer time. Better sorting techniques need to be devised, and this is largely a technical design problem. Perhaps special purpose machines restricted to sorting and equipped with only enough hardware to accomplish it, are the answer. This may be a matter of economics and perhaps here again we in business have failed to give you a clear specification of how much sorting we have and how fast it needs to be done. If you need these data, they are much more readily available than the data on random access.

Equipment Reliability

Another area where work needs to be done is the area of equipment reliability. This divides into two parts—availability of equipment and reliability of results. In the availability problem the maintenance engineer and the businessman are quite far apart. No businessman wants equipment that he has to shut down frequently and yet the maintenance engineer knows the need for constant overhaul. Perhaps there is some middle ground upon which they can meet. Most businessmen are willing to concede that maintenance is necessary but will insist that unless they get 95 per cent useful time out of a machine it has a reduced value to them. To have to give up a machine for an hour or more every day or every shift just for preventive maintenance in addition to actual breakdown, seems to him to be a hardship. You may point out that this is an economic problem and better machines can be built if we are willing to pay for them. This may be true and it may eventually come to that. Just at the moment, however, the disruption to our carefully timed business processes alone makes maintenance a major factor in the decision to use an electronic computer and I feel that work should be done to provide more available time with presently available equipment. If it cannot be improved, then you need to tell us how much more it will cost to bring it up to our standards. We can then decide whether or not we are willing to spend it or accept the alternative of reduced usefulness.

The second area of equipment reliability has to do with the reliability of results. Here we encounter one of the basic differences in the philosophical concept of computers. One school holds for internal hardware duplication sufficient to assure practically perfect operation while the other depends on less hardware and more programmed checks. Which is right only time will tell and after we businessmen have more experience with each we will be better able to judge their merits for ourselves. But I think you need to do some work in this area too. This is not only a sales pitch. There must be some facts to work on. Surely you must be building up a body of data on the pros and cons of this question. Can't you computer designers reconcile these two conflicting camps and give us at least a theoretical answer? Perhaps your answer will be that there is no one best way and each has so much to be said for it that there is a place for both concepts. Just at the moment, however, the businessman finds himself between two fires and has difficulty deciding between them because each sounds so logical when he hears it advocated by its exponents.

Special vs General Purpose Equipment

The last area needing consideration that I want to discuss with you for a minute is this question of special vs general purpose equipment. Here is an area where there is still a wide difference of opinion. Most of today's equipment is general purpose, probably for historical reasons, yet we do begin to see some signs of the development of special purpose equipment. This will be useful in specific cases but it is yet to be determined if it is the final answer. Business changes so rapidly and so much that equipment restricted to today's needs may not be at all usable next month or next year. If you can design and build special purpose equipment considerably cheaper than general purpose equipment, perhaps we can afford to use it while it fits our purpose and then discard it. If you cannot do so because of development costs and lack of volume production, we shall have to continue to use the larger general purpose equipment.

You are probably aware that this problem is giving many persons considerable concern, but you may not be aware that the electric and gas industries are trying to do something about it. Jointly they have engaged the computation laboratory of one of our large universities to carry on a research project to determine whether or not special purpose equipment is practical for the utility industry, and if so, the nature of it. I understand some of the banks have done about the same thing. What the result will be is of course uncertain. But I believe you should weigh carefully with us the question of whether or not we need all the things you provide now and what would be the savings by eliminating some of them for certain applications. Again perhaps we need to study this as much as you do because we may not have outlined our needs to you accurately and clearly enough to enable you to visualize what they represent in hardware.

If you will review my remarks, particularly those dealing with areas to be studied, you will be struck as I was by my frequent reference to the need for closer coordination between the designer and the user of the equipment. This is the note on which I should like to close. You have produced some marvelous pieces of equipment and we in business are grateful for them. But in the evolution of any new field there comes a time when basic invention must give way to refinement and improvement aimed at greater usability. I think we are very close to that point today. This can only be accomplished by a careful study of the needs of the user and the development of an awareness of the manner in which your designs fall short of maximum convenience. I do not know the mechanics of accomplishing this either on the design side or the user side but I believe that it is the next major step which needs to be taken if the computer industry is to flourish and we in business are to have the best machines for our problems. I sincerely hope that the immediate future will see the development of this close cooperation. If my thoughts and remarks today act as a stimulant toward that end I shall feel that they have been worthwhile

Discussion

W. K. Halstead (RCA): What steps are you taking for interchange of data between the Univac's and the 705?

Mr. Porter: We are not taking any steps for that purpose. We believe that we have sufficient volume to justify machines in both areas. The IBM Company in New York City has a service bureau, and Remington Rand has a service bureau, both within a few blocks of our office. We feel there is sufficient backup in that area to support us. We have sufficient assurance from the manufacturers that they will assist us if necessary.

R. S. Gillett (National Security Agency): In a multiple activity application would it not be better to draw programmers from each activity; or, in other words, let each activity program its own problems?

Mr. Porter: I think it would be. Actually, that is one of the things we have done in building up our staff in the general and customer accounting area; revenue accounting, as you know it. We have drawn people from the sales, auditing, purchasing, and engineering departments. Our thinking there is not only to bring the viewpoints of those departments into the problem, but, when the job is completed, to allow those people to go back to their own departments and view the problems of their particular departments to see if some of the spare time of the computer could be used.

E. L. Harder (Westinghouse): What job will Consolidated Edison put on the computers on order? What is the expected payoff on the computer use?

Mr. Porter: The Univac will take on customer accounting which includes everything from the time the meter reading is brought in to the time the bill is calculated and includes the processing of collection follow-up, statistical and rate analysis, and similar data. The work on the 705 would be payroll to be followed probably by store accounting and the other areas such as accounts receivable and general accounting.

On the payoff until we have completed our installation we cannot be sure but certainly think that within five years we will be pretty well in the clear.

Operations Control with an Electronic Computer

B. F. BUTLER†

I CONSIDER it a great privilege to participate in the discussions of this conference. I can think of no more timely panel topic than the Role of Computers in Business. We are all familiar with the enormous impact that computers are having on industry today, and the promising future for computers that lies directly ahead. I believe we should face this dynamic situation with full realization of both the potentialities and the difficulties that confront us.

The role that electronic data processing equipment is expected to play in the national economy is reflected by the capital investment in such equipment. Electronic data processing machines installed now, plus those committed for delivery during the next two years, represent a capital investment of over one billion dollars. This mammoth investment is in equipment that may be obsolete in five to seven years. This means that a return on the capital investment of about 20 per cent annually must be obtained in order to break even at the estimated obsolescence. Twenty per cent of one billion dollars is $200,000,000 annually. In other words, the benefit to the national economy must be approximately $200,000,000 annually in order to break even on this billion dollar capital investment.

In terms of the individual data processing machine, the business manager should be able to recover his capital investment in five to seven years and obtain a reasonable rate of interest on his capital. If the equipment is procured on a rental basis these considerations are factored in the rental fee. In any case, when the business manager invests in an electronic data processing machine, either by purchase or by monthly rental, he has incurred a major expenditure that must be compensated for in operating benefits.

I cannot help feeling that many of the decisions to buy or rent electronic data processing machines are impetuous acts on the part of some executives who feel intuitively that the large computer will pay big dividends. This brashness is based upon the deep seated belief that electronic data processing machines *per se* will reduce paper work enormously. There seems to be little realization that this billion dollars worth of machines can generate paper in quantities never before possible. The executive may find himself creating a flood of paper from a battery of high speed printers each grinding away incessantly, like the sorcerer's apprentice who summoned the broom to carry water but could not stem the flood.

I am reminded at this point of the supervisor of a tabulating installation I visited a few years ago. He measured his productivity by the miles of tabulating paper he consumed. With great pride he told me that during the previous month he had supplied the accountants in his company with 700 miles of tabulating paper. Just think what that ambitious fellow could do with a battery of present day high speed printers. Whether electronic data processing machines reduce paper work or become paper generating monsters depends upon the quality of the planning and thinking through process prior to using the machine. Unless superlative plans are evolved, the paper generating characteristics of the machine will dominate.

The quality of the data processing planning depends upon the competence and versatility of the scientists and procedure analysts responsible for the systems work. I say "scientists and procedure analysts" because I believe such a combination of talents is a necessary prerequisite to effective data processing. And I say "prerequisite" because I believe the systems work should be done prior to procurement and installation of the data processing machine.

Usually we find procedure analysts on the systems team. We do not seem to err in this direction. On the other hand we are all familiar with installations where untold difficulties were encountered because the systems team lacked the necessary scientific components.

Now that I have belabored the need for an adequate return on capital investment, and the need for a team of scientists and procedure analysts to do the systems work, let us turn to one of the basic principles essential to successful operations control.

Let us consider a very simple type of operations control, the traffic light on a street intersection. I think we can all agree that this is an effective control system. Yet, while it is doing its controlling, the traffic light does not do many of the things we normally associate with operations control. For example, it takes no measurements, it performs no analysis, there is no feed back and there is no paper generated. The green and red lights merely go on and off in a predetermined manner. What is it then that makes the traffic light an effective means of control? It is the traffic rule that the driver of the vehicle shall go when the light is green and shall stop when the light is red. Rules of this type are frequently referred to as decision rules. And it is the presence of effective decision rules that make for efficient operations control. Frequently, however, decision rules are ill defined or entirely lacking. When this is true I think we can say quite dogmatically that effective operations control is unlikely.

In its simplest form the decision rule tells us when corrective action is necessary to maintain control. Statistical quality control charts are a good example of this type of operations control. When the measurements fall within the confines of the control chart limits the decision rule says that no investigation is warranted. When a measurement falls outside the control limits on the

† Hanford Atomic Products Operation, General Electric Co., Richland, Wash.

chart the decision rule says that the process is probably out of control and an investigation is warranted to locate and remove the assignable cause.

Budgetary control is another example of operations control where the simplest form of decision rule is used. Here actual expenditures are compared with pre-established budgets. If we are within budget limits the decision rule says that no investigation is called for; if we exceed the budget the decision rule calls for an investigation of the causes.

Note that this is a very elementary type of control. The measurements and decision rule merely tell us whether to consider the operation in or out of control. It tells us whether or not to take corrective action. It does not however, tell us what corrective action to take. Further analysis of the data is required to determine the corrective action necessary; or additional data or experimentation may be required to locate and remove the cause of the difficulty.

In the next level of sophistication in operations control the decision rule not only serves to indicate *when* corrective action is required, but also stipulates *what* corrective action to take. It is this level of control that should be attained with general purpose electronic data processing machines. In many situations all the important factors that enter into a routine decision are known. In such cases the data processing machine can be programmed to analyze these factors and record the resulting decision.

The highest level of operations control is attained when the decision rule is fed back automatically to the operation to obtain the necessary corrective action. In this case we have automation. It is here where the special purpose computer excels.

So much for the importance of decision rules in operations control. I think I have made the point I wish to stress, namely, that effective decision rules are necessary for efficient operations control.

Let us turn our attention to Fig. 1 which illustrates graphically the relationship between planning, decision rules, measurements, and analysis in operations control. Note that two closed loops are shown.

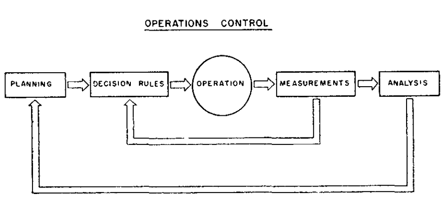


Fig. 1

Let us consider the planning function first. The use of decision rules requires that we think ahead. We must program into our data processing system each decision that will be reached under each set of conditions that might occur in the future. This calls for a different kind of thinking than we are normally used to. Ordinarily we reach a decision after we review the data. The decision rule requires that we plan before the data is collected what our decisions will be under various sets of conditions that might occur in the future.

The plans that are designed to control future operations are executed by the inner loop shown in Fig. 1. Appropriate measurements from the operation are fed to the decision function to obtain control instructions that are fed back to the operation.

It is not always possible, or even desirable, to have decision rules that will take care of all contingencies. There will be circumstances when a careful analysis of the situation is preferable to a precast decision. In such cases the outer loop in Fig. 1 should be called into play. Here the measurements from the operation are analyzed and special plans are evolved for handling the control situation. The outer loop also serves to keep the inner loop in adjustment, or to change the characterstics of the inner loop to take care of changing conditions.

It is in the outer loop where operations research comes into play, specifically in the box marked analysis. As Mr. J. S. Coleman has mentioned, management is interested in knowing what alternative solutions there are to business problems. Unlike engineering problems there is rarely one "best" solution to a complex business problem. The most we can hope for is a number of alternatives each with advantages and disadvantages. As Mr. Coleman mentioned, it is here where techniques like linear programming can be helpful.

We have used linear programming models up to 150 order in size. We have found that linear programming does not give a detailed schedule by which the entire operation can be run. The power of this technique lies in the ability to vary the conditions, capabilities, requirements, and value criteria that surround the operation, and to obtain the consequences of such changes. As Mr. Coleman indicated, it is a tool for exploring operating alternatives.

It is interesting to note that at first our operating people did not accept the results given by this model as valuable, because detailed supporting schedules were not prepared as "back-up" data. Increased understanding, however, has led to much greater confidence in the model and the results of manipulating it. Equally important is the fact that the attempt to understand this model has shown operating people that there are certain concepts, or principles, in their operation that allow them to understand the important aspects of the operation without complete dependence on intuition on the one hand, and massive detail on the other.

Let us consider a simple example of operations control so that we may better understand the implications of Fig. 1. We shall assume that the operation to be controlled is the central heating system in your home. If the thermostat instructs the furnace to turn on at 60° and off at 65° this is our decision rule. The inner loop on Fig. 1 executes this planned control. The temperature is measured and the thermostat sends instruction to the furnace to turn on and off. However, if the housewife feels cold she is performing the analysis indicated in the outer loop. If she decides to raise the temperature she is now performing the planning function. If she adjusts the thermostat to turn on at 70° and off at 75° she has changed the decision rule to conform with her analysis and planning.

The inner and outer loops shown in Fig. 1 can be utilized in operations control with or without an electronic computer. The example of the central heating system illustrates this. Nevertheless, it is the advent of electronic methods of control and computation that has increased the potentialities of operations control by a very large factor. I feel that we are justified in treating operations control with an electronic computer as a major event in industrial evolution. And I think there is every evidence that future historians will treat it as such. Surely it is not difficult to foresee a bright future for electronic controls. But it is the transition period, the period of electronic adolescence we are now entering, that we must use discrimination with our onrushing billion dollars worth of electronic data processing machines. Unless our competence matches our brashness we may become the victims of our intended benefactors.

And now I have come to the end of my story. The point I have been attempting to stress is that there is a right way and a wrong way to go about operations control. The right way includes the use of effective decision rules. The wrong way entails reserving decision until a review of the control data has taken place. This usually means that reams of paper are reviewed in an attempt to reach a decision. The decision is delayed, and the corrective action may be late.

In business we are about to face a serious shortage of personnel due to the low birth rate during the depression. This will have the effect of increasing the role that electronic computers will play in business. Many important tasks will remain undone unless the machines do them for us. This will be true not only in the factory, but in the office and scientific laboratories as well. Electronic computers will be called upon to obtain greater productivity from the limited manpower resources.

And so we have before us the challenge of the business world to fashion electronic data processing into a powerful managerial tool capable of assuring an accelerated prosperity.

Discussion

W. M. Young (Pepperell Mfg. Co.): There are two kinds of obsolescence—engineering and economic. When you gave a figure of from five to seven years for obsolescence which kind did you mean and how did you arrive at your figures?

Mr. Butler: That is quite a question. I think I was referring to both engineering and economic obsolescence. It may be possible that some of the larger machines used by large companies can be hand-me-downs to small companies and perhaps this might be an economical thing for the small company to do.

Talking in terms of the business manager who is making an investment I think it is the engineering and scientific development that will take place that will make his machine economically obsolete in from five to seven years. I think if he doesn't take the new machines in that time his competitors will and it will be an economic factor.

J. A. Sperling (Western Union): From your experience on operations research can you give us an example of how you have seen it used in practice?

Mr. Butler: Well, we have been conducting operations research at the Hanford Atomic Products Operations now for about two years. We have constructed certain mathematical models which represent the operations. These models usually run from somewhere between 100 to 150 equations. With these models we have been able to develop numerous operating characteristics that have been extremely beneficial.

For example, on the very first run that we made, using this model, the model questioned the reopening of an old plant that had been shut down. Millions of dollars would have to be appropriated to reopen this plant and operate it. Our operations research indicated that Hanford could be operated without the use of this plant. This called for a considerable reversal of plans as the result of our very first run. Although no action was taken as a result of this first run, management subsequently decided against the reopening of the plant. It is difficult to tell exactly how much influence operations research had in this decision. Undoubtedly it has some influence.

Figures on capital investment in equipment were estimated from the number of machines now in operation plus the number of machines that have been ordered for delivery. The obsolescence figures are an opinion.

Now, other benefits that have come from this same operations research have been in the area of giving management a better understanding of the operations, a different way of looking at it. We had always looked upon the Hanford Works as the plant for manufacturing plutonium. However, when we looked upon it as a uranium processing plant we got a much better understanding for planning purposes. This is pretty sketchy but it will give you an idea of what we are attempting to do.

We not only manufacture plutonium but we operate the village of Richland, a transit system and some hundred miles of railroad. We made a study of the transportation system and presented to management some 42 alternative ways that they could operate this business. We were able to show the tangible cost savings of each alternative from zero up to a half million dollars a year. However, as the savings went up the intangible deterrents in the form of employee and public relations went up. At first management felt they could do nothing with the results of this program, but later adopted one of the alternatives. This gave some measure of the value management was placing on employee-management relations.

The Place of the Special Purpose Electronic Data Processing Systems

R. E. SPRAGUE†

IN RECENT literature and discussions the words *Special Purpose Computer* have been appearing with greater frequency, for example, in an editorial by John Diebold in August, 1955 *Automatic Control* entitled "There's A Need For Special Purpose Computers." At a recent meeting at Stanford Research Institute, eminent business representatives reached the conclusion that special purpose systems should be given more attention in business and industrial applications of "electronics."

A careful analysis of the reasons behind this apparent trend and of the place of the Special Purpose Systems would seem to be in order.

History

To fully understand the status of special purpose machines, it is necessary to trace the history of development of the so-called "general purpose computers" otherwise known as "internally programmed data processors." Unfortunately, it is also necessary to outline some of their weaknesses. That this has not been done to any great extent in the past is partly because of a lack of enough knowledge in the applications end of the "electronics" business, and partly because the majority of concentration in the field has been focussed on the use of programmed machines.

Many lay persons have stated that the first computers were designed to solve scientific and engineering problems and that this is the principal reason they have been so slowly applied to business problems.

As a former general purpose computer designer starting in the early days of the field, I can say that this is not entirely true.

Türing System

As early as 1939 a British mathematician named Türing proposed a theoretical system which would solve any problem that could be presented in logical terms. The system had properties possessed by the considerably later-developed internally programmed computers.

The first internally programmed computer designers essentially developed Türing's ideas. They realized that it was possible with available electronic components to build machines which would solve any problem or carry out any sort of data processing desired whether it be a scientific or business application.

The concentration of design, however, was first on the mechanizing of Türing's principles.

† Manager, Field Operations, Teleregister Corp., Stamford, Conn.

Thus the intention of the early designers was not to build a machine to solve scientific problems, but rather to produce a machine that would, following certain theoretical principals, "do anything."

To understand one of the reasons why the machines were applied slowly to business applications, it is necessary to explain these principles. The most fundamental of the ideas is that all logical data processing operations, be they scientific or commercial, can be constructed from a reasonably small set of very basic logical data manipulations (from this point on referred to as instructions). Some of these are:

a) Storing data in a memory of some kind;
b) Adding;
c) Subtracting;
d) Multiplying;
e) Dividing;
f) Shifting, inserting and extracting data to and from other data;
g) Comparing relative status (or order) of pieces of data;
h) Moving data from one location in storage to another.

The internally programmed machines were designed with a memory or storage medium to hold data and with the ability to carry out sequences of these extremely *basic* instructions.

The sequences themselves (from this point on, referred to as programs) were stored in the same "memory" as the data and very complicated operations, and hence any kind of data processing could be performed.

Decision instructions built into the machines enabled them to perform alternate programs or to repeat a program depending on prior results or on predetermined standards.

Programming Time

As attempts were made to apply these machines to business problems it became apparent over a long period of time that programming for business applications was a much longer, more tedious task than anyone had imagined. This was not as true for scientific problems although, even in that field, programming and program debugging[1] times were underestimated in the beginning.

[1] The job of checking out a program in actual operation for errors and omissions.

Input-Output

To be fair it should also be mentioned that development of reliable means for rapidly getting data into and out of the machines lagged considerably behind the machines themselves. Business data processing in general involves a much higher percentage of processing time devoted to inputting and outputting. However, it can be safely stated that this problem has been reasonably solved in the last few years and that the programming problem has not.

Alphabet Analogy

To see what led to the present situation in programming, an interesting analogy can be drawn. The first internally programmed computer designers theorized, in effect, that by building machines which could carry out a universal set of very basic instructions, it should be possible to make them do almost anything.

This is analogous to saying, "By creating a small set of characters known as the alphabet, it is possible to create the English language."

That it can be done, no one doubts; but the time and effort required to do it could very easily be underestimated.

In order to create the English language from the alphabet, it would first be necessary to make up words, then sentences, then paragraphs, etc.

In a similar sense, to create the programs to enable an internally programmed computer to accomplish any type of business data processing, it has been necessary to build up larger and larger spheres of operations from the basic set of instructions. This is what takes all of the man hours.

I make the analogy of the alphabet to point out that the idea of building upon instructions basic enough that the machine can do anything means using *extremely basic instructions*. They are so basic that for any given program it takes very many of them and a great deal of programming attention to the minute details of the over-all operation.

There are indications that the General Purpose Computer manufacturers recognize this problem. Several machines have had less basic instructions (such as sort and merge) added to their repertoire.

Today, little doubt exists that programming times and costs are high. For example, it has taken, in one known case, many man years to complete just the first of several programs intended for a machine being applied to commercial work. So far it appears that programming costs will approach the cost of the computer itself.

One of the factors increasing programming costs are the wages now being drawn by skilled programmers. They are a very scarce commodity. After a businessman's own personnel become skilled in programming, (an expensive training is involved), they become very hard to hold. Turnover is high. This situation, by all educational indications, will continue for several years.

Special Purpose Systems

Special purpose systems which might better be described as tailor-made or custom designed systems get around the programming problem by being constructed out of much less basic operations (by analogy, sentences instead of letters) which are all selected to fit the exact requirements of the application. These operations are built-in rather than programmed. If any variations take place in the sequences of these macro (rather than micro) steps, they are controlled automatically by input devices.

No programming is necessary. No programmers (highly paid compared to operators) are necessary. No programming training is necessary. The machine operators can be easily trained and can be nonskilled. The devices they operate are quite familiar to them and use language on the keyboards that has always been in use in their business.

The user of the special purpose system needs only to concern himself with renting or purchasing the equipment and a small amount of training of operators he usually already has employed. Naturally the systems study necessary for any electronic mechanization must also be performed in advance.

Adding Applications

Special purpose systems have other *general* advantages. Arguments have been advanced that if a special purpose system user desires to add another application or to expand his initial application, it is harder or more expensive to do this with a special purpose than with a general purpose system.

It is true that more equipment would have to be added but it is also true that properly designed special purpose systems can easily be expanded or added to.

As for cost, if the cost of programming for a new application approaches the cost of the general purpose equipment, as has been previously pointed out, then the *total* cost of the *final* general purpose system with programming may well equal the *total* cost of the final special purpose system.

Delivery

Another argument advanced in favor of the general purpose system is the time delay involved to acquire the additional special purpose equipment as opposed to buying the general purpose machine at the outset with no further delays. Actually, such reasoning begins to make less and less sense as businessmen become more aware of the time required to prepare, check out, and put into operation a new program. Times far in excess of those originally estimated have extended programming periods to as much as one year and longer. The delivery schedules for special purpose equipment are in the same range.

Size of Application

General purpose computers come in certain standard sizes. If a business has an application just a little too large for one size or smaller than the smallest general purpose computer capacity, then the business is paying for more than it uses until additional applications are programmed. Special purpose systems can be made to tackle any size application with no more built into the equipment than is actually required.

Mechanizing by Degrees

The special purpose system has a psychological and, in some cases, practical advantage over general purpose systems. This is the possibility of "walking before you run" by building equipment and training operating personnel on only one application or part of one at a time.

This has the double advantage in the special purpose system of saving money in equipment costs and programming and also the attendant reduction in the amount of disruption of the organization. Special purpose input and output equipment and input languages can be designed to match the procedures already in practice. Semi-manual operating procedures can be used at first giving management a more secure "feeling" that the entire electronic system is not out of their control.

While it is possible to walk before you run with the general purpose system, there is a very great tendency to go "whole hog" in mechanizing an application. When this is done, management has the feeling that they can't reach into their data processing system, that they have lost touch with it.

More changes in operating procedures are necessary because input and output devices used with general purpose computers are standardized. Many "still required" manual operations have to be changed and forced to meet the rigidity of the general purpose input and output system. Codes must be learned or standard keyboards used, etc. Thus hidden transition costs must be added to general purpose equipment and programming costs.

Scheduled Systems

In listing these general advantages, I do not mean to imply that a custom-designed system with a wired-in program can be economically used in all business applications now being programmed for General Purpose Computers. In fact, credit for the success of special purpose systems in some types of applications can be attributed to a different factor. General purpose computers are regimented (scheduled) devices with but a single input and single output. Generally speaking they perform one operation at a time in serial fashion.

While there are exceptions here and there, the machines patterned after Türing's superb logic cannot add while they subtract, or print out while they read in, or multiply while they print out, etc. This is because they are all designed with a single control unit which regulates all operations in a serial fashion. Again, speaking generally, to make a general purpose machine do more than one thing at a time would require more than one control unit. This would double that section of the machine and introduce crosstalk problems if the control units were sharing the same memory. Also, some definition must be made as to which control unit takes precedence, implying a super control unit.

Some progress has been made in the direction of parallel operations[2] by using buffer registers, time sharing and building separate units for special operations such as sorting. However, the general purpose machines are still fairly limited compared to special purpose systems in this respect.

Special Applications

Certain types of "on line" applications just naturally generate requirements which do not fit these regimented general purpose computer characteristics. Some of these requirements are:

1) Multiple simultaneous inputs.
2) Multiple simultaneous outputs.
3) Output answers both printed and visual at input points which originated requests for the answers.
4) Simultaneous inputs from many remote locations with immediate answers back necessary.
5) Necessity of interrupting regular data processing flow to handle inquiries.
6) Necessity of processing data on a schedule regulated by some random control such as customer calls.

All of these requirements have been met very successfully by special purpose systems over the past few years.[3] Another type of application not efficiently handled by General Purpose Computers is that involving the physical handling of paper as well as information processing. Special purpose systems which look promising are under development for these applications.

Automatic Programming

One significant activity called automatic programming may reduce total general purpose system costs. The basic idea in this activity is that the computer should do part of the programming job thus cutting time and personnel costs.

However, the idea applied to business problems has not proven too successful as yet. There are too many variables and differences between one business application and another to allow much standardization of

[2] Notably on the UNIVAC, BIZMAC, ELECOM 125.

[3] Examples: Teleregister's Reservisor at American Airlines and Teleregister's Stock Exchange Bid Asked System at Toronto Stock Exchange.

so-called sub-programs. It is possible to prepare very generalized automatic programs for a given type of business application, such as payroll, but only at the expense of a good deal of computer time. The additional time is required for the computer to delve into the many possible variations between one company's payroll and another and to decide which of these variations should be satisfied.

Machine Costs

On the larger computers, the cost of machine time devoted to automatic programming such as compiling and generating routines may amount to an appreciable portion of the cost of personnel programming for the specific application. This is true because machine time may be worth several hundred dollars per hour while a programmer's time is worth perhaps three to six dollars per hour. Thus the savings in programming costs will not be as high as one might expect and they will be partially canceled out by higher processing costs.

Conclusion

The Special Purpose Data Processing System appears to be taking its place in many applications which have requirements that do not match the characteristics of general purpose computers. Also, as more and more businessmen, and for that matter, technical people realize the exceedingly long times and high costs involved in programming, they may turn to the possibility of special purpose systems.

Discussion

E. L. Harder (Westinghouse): Please give some examples of the special purpose computers that you are talking about; that is, examples of specific applications being considered.

Mr. Sprague: It is rather hard to dig out any clear-cut examples that can be said to have proved themselves economically. Of course, I am familiar with one and that is American Airlines at LaGuardia Field. I am sure most of you have seen enough literature on that so I don't have to describe it but it has proved itself. The Toronto Stock Exchange bid-asked storage system which I mentioned earlier is another example.

Another type of machine that is appearing now is that represented by the Remington Rand file computer. It can be used with a number of input and output devices that make it adaptable to special purpose applications.

Teleregister is also working on other types of special purpose equipment which will make appearances within the next year; primarily on inventory control and transportation systems.

R. A. Butterworth (General Electric): "With special purpose equipment, no programming, or high-priced programmers, are required." Who, then, specifies the characteristics of, and designs, the special purpose equipment? Also, what happens to the special purpose equipment when a change in procedure is indicated by, say, operations research?

Mr. Sprague: The specifications are laid down by the customer operating personnel and the special purpose manufacturer. Up to a certain point the systems studies performed by both the manufacturer and customer are very similar to those of a general purpose system. Beyond that point, rather than enter a programming phase—and this includes all those jobs that require instruction—the engineers of the manufacturer enter into a computer design stage. A certain amount of money that appears in the final price, or rental price, of a special purpose system represents this engineering. When a change occurs from the original design, that change could be entered if realized soon enough by the manufacturer. Of course, if the equipment has been completed it would be necessary to expand or alter it in some way. This could be done if the deisgn has some flexibility for making changes. Actually, in one of Teleregister's present equipments expansion possibilities have been allowed to take care of the possibility of double the present volume of storage as well as information traffic.

J. H. Burns (Liberty Mutual Insurance Co.): Did I understand you to imply that special purpose computers require less systems research than general purpose computers? Could you explain?

Mr. Sprague: No, I don't think I meant to imply that, other than to state that the certain areas in the system study area, let us say, associated with input and output equipment and changes in operating procedure, would not have to be affected immediately; particularly if the input-output equipment takes present procedures into account.

To give you an example, rather than use a standard typewriter keyboard for entering information on reservations, one designs a special keyset which uses input systems familiar to the operating personnel of a reservations bureau.

L. C. Hobbs (Sperry Rand): With special purpose equipment are you not exchanging logical designers for programmers? If the problem remained the same for special purpose computers then canned programs could be used for general purpose machines?

Mr. Sprague: Yes, you are to some extent exchanging logical designers for programmers. The question of whether canned programs can be used more economically than logically designed elements in a machine, centers around the question of how elementary the commands are as opposed to built-in operations in a special purpose machine. Taking any one application, it seems to me that the number of program steps can be greatly reduced when wired into a special purpose system, because of the use of macro (rather than micro) steps. The simpler the application, the greater will be the advantage gained by "wired in" programs.

Electronics in Financial Accounting

B. J. BENNETT†, K. R. ELDREDGE†, T. H. MORRIN†, J. D. NOE†, AND O. W. WHITBY†

Elements of Financial Accounting

THERE ARE FOUR essential elements in financial or dollar accounting. One of these elements is represented by the document of authorization. Another is a set of records for legal purposes. A third is a set of records for analysis, and a fourth is a set of records for control.

In any accounting system for management of money, documents of authorization are essential and these documents, bearing the signature of the authorizing agent, must be preserved for legal purposes and for control of the system. The set of records required for legal purposes must be available to settle questions of equity, since ownership of various portions of the money being handled must be clear at all times. The records for analysis must show the amount of money that is available at any time. They must also show the dollar amounts which are tied up by the processing method and which, hence, are not available for other uses. The records for control are needed for control of accuracy of the accounting procedure and as a basis for management decisions concerning funds.

This discussion will concentrate on bank accounting for commercial checks. It will not stress general principles, however the problems, conditions, and solutions for commercial check accounting do apply to many dollar accounting problems. Commercial check accounting involves all of the above elements of financial accounting systems, and it has been the subject of a five-year study by Stanford Research Institute for the Bank of America—a study that has resulted in an integrated system to handle commercial check accounting.

This discussion will attempt to point out the principal factors which affect design of a system to handle commercial checking, to show the consequences of these factors, and to describe, briefly, a system developed under their influence.

Before proceeding with a description of the problems, let us first examine the operations that take place in the handling of commercial checks.

Work Flow in National Banking

The actual work flow within the national banking system is very complex. It would be quite simple, although quite pointless, if each customer who wrote a check presented that check only at the window of the bank at which his account is held. This is far from the case. As a national average, only one out of five checks drawn on a given bank is charged immediately to an account in that bank. The other four must be returned through the national network of other banks.

To go into details of all possible means by which checks and deposit slips travel from bank to bank, would be an impossible task for this brief paper. However, certain concepts are important, and simplified diagrams can be used to establish these concepts. It is useful to define "clearing items," "transit items," "go-backs," "float," and the so-called "on us" items. It is also illuminating to note the volume of work and the pressure of time involved in the banking system.

Fig. 1 illustrates a clearing arrangement. This figure is very simplified and shows flow in only one direction, *i.e.*, from Bank A to Banks B and C. In looking at this figure, bear in mind that two other flow arrangements should be superimposed upon it, *i.e.*, from Bank B to Banks A and C, and from Bank C to Banks A and B.

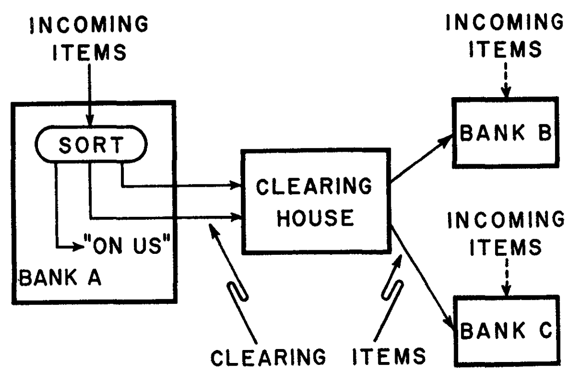


Fig. 1—Clearing arrangement between competitive banks in same area.

The problem is this: Each bank receives over the teller's window, or through the mail, a mixture of items belonging to its own accounts and items belonging to other banks. These must be sorted out and exchanged. This commonly takes place through a clearing house which is financially supported by all the banks involved. These competitive banks agree to exchange items at a specified hour each day and each of them keeps a record of exchange in its so-called general ledger accounts. Bear in mind that clearing arrangements deal with banks located in the same geographical area. Also note the "on us" items which are charged to accounts in Bank A. This represents a great volume of work.

Fig. 2 illustrates some of the ways in which the same problem is handled when the banks are in different geographical areas. Three areas are shown, each having a group of banks with local clearing arrangements and

† Stanford Research Institute, Menlo Park, Calif.

showing the flow of transit items, *i.e.*, those which are presented at one bank but belong in another. In this case, many combinations are possible and Fig. 2 shows only three kinds of situations. In Area 1 is shown one bank of a system composed of many banks or branches controlled by a head office. Area 2 shows a single large bank with no other branches but with clearing arrangements with local competitors. Area 3 indicates two smaller banks, each with its own clearing agreements, but for which the Federal Reserve Bank performs the function, in some respects, of a head office. Again this figure shows flow in only one direction and there is actually flow in all directions. This is a very simplified picture—in actual practice there are many more arrangements.

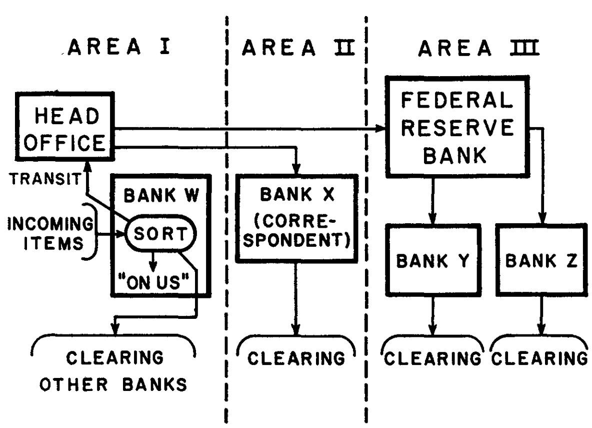


Fig. 2—Transit arrangements.

Perhaps these comments on clearing and transit give a hint of the complex network which exists. Now add to this the figures on volume of work handled in the system. In 1939, approximately $3\frac{1}{2}$ billion checks were written. During 1952, this rose to 8 billion and—if the trend continues—1960 will see 14 billion, and 1970, 22 billion checks written. Everybody hastens to point out that this trend may not continue at such a pace but no one argues with the fact that the volume is growing. Another factor of extreme importance has emerged from national studies. Each check is handled an average of 6 to 8 times during its progress through the banking system. By "handled" is meant the actual entry of the dollar amount of the check through a keyboard.

To the complexity of the work flow and the volume of data handled, one must add a time element. During the process of all this exchanging of checks, a certain number of "go-backs" occur. These are checks which cause overdraft or are drawn against uncollected funds, checks which are postdated, have stop payments against them or have irregular signatures, and checks which are missorted. These "go-backs" must be detected rapidly by the receiver and returned to the sending bank. The amount of time which can elapse before "go-backs" can be refused by the transmitting bank is rigidly fixed by agreements between banks and varies, depending upon geographical area, between approximately 3 hours and 48 hours.

The concept of "float" is also important. Brief mention was made, in the opening sentences of this discussion, to funds which were "tied-up" by the processing method: As an example, suppose Mr. Smith deposits in his bank in Boston, a check for $1,000 given to him by a man who has an account in San Francisco. Then suppose Mr. Smith immediately writes checks which use up half of that $1,000 deposit. If it takes three days for the Boston bank to collect from San Francisco, the Boston bank has loaned Mr. Smith $500 for those three days. There are many more ways in which "float" can occur, but this should illustrate the idea.

From all this emerges a very complicated system problem. A large number of individual pieces of paper travel through a very complex network. The present manual system demands handling them over and over again, each time incurring the possibility of human error, and a rigid time limit surrounds the entire operation. Out of this system, two distinct but closely related problems emerge. One is the transit operation. The other is the problem of keeping track of the "on us" items, *i.e.*, the bookkeeping on the accounts held within each bank.

The five-year program which the Institute carried out for the Bank of America dealt with the bookkeeping and clearing for its branch banks. This was chosen because it is a large volume job, it is one of extreme importance, it needs speeding up to reduce float, and it primarily affects internal procedures. The transit problem immediately involves other banks, their procedures, and their standards. Out of this five-year program, techniques have evolved for the bookkeeping and check handling work and extensions in techniques are being developed which are applicable to the transit problem.

Factors Affecting System

Now, let's look at some of the main factors in this job and consider them from the standpoint of how they affect system design. A few of these factors seem unique to the banking business, but counterparts can be found in most applications which involve dollar accounting. Among those which are most important are the following: the need for economy, the volume and nature of the source documents, the high degree of accuracy required, the fact that system changes must be evolutionary, not revolutionary, and the nature of the filing problem required to keep track of the information.

Net economy is essential, of course, and this must include all factors—operating cost, equipment amortization, and the effect of the national tax structure. As of this time the computer industry must face fairly short equipment write-off periods, since the average user cannot help looking upon the venture as experimental, and because he has a fear of obsolescence. Hence equipment write-off is, today, a big factor.

The need for economy has one important influence on the handling of commercial checks. It dictates centralizing a large number of accounts within a given machine and, hence, implies central bookkeeping. This is largely because control circuits required to handle all operations represent an investment so great that it is necessary to spread this investment by handling a maximum number of accounts, thus decreasing the cost per account.

The time schedule is very rigid. Even though the volume of work may vary a great deal from day to day, it is difficult in bank accounting to carry over part of one day's work into the next in order to reduce the peak load and hence reduce the maximum rate at which the system must assimilate and process data. As mentioned previously, this rigid time schedule arises because of the clearing arrangements and the treatment of "go-backs." However, another important point affecting the time schedule arises from the fact that it is necessary to "prove" the day's work to find errors just as soon as possible after the receipt of the last item for the day. Furthermore, it is not practical to hold an entire day's work, then to process it at high speed after the last item has been received, since this would greatly increase system cost and complexity. Hence, it is necessary during the day to do all the work possible so that when the last item is received, most of the actual posting to individiaul accounts and altering of current balances is finished so that the system can proceed directly with a recapitulation and balancing.

This combination of time schedule and large volume of information forces one to consider simultaneous operation of the input, sorting, posting, and printing functions.

It is also necessary to have means for a rapid diagnosis and correction of errors so that the entire system can be balanced quickly at the end of the day. The tight time schedule obviously calls for minimization of down-time. Therefore, careful attention must be given to maintenance procedures.

The third principal factor affecting system design, *i.e.*, the volume and nature of the documents, raises some very difficult problems.

The large number of checks and deposit slips which must be handled demand high speed processing and transcription. While it is possible to read the information from these documents when they are returned to the bank so that this information can be processed, it is still necessary to sort, examine, and file the checks and return them to the customer. Therefore, any system designed to deal with the commercial checking account business cannot ignore the paper-handling and transcription problems. Furthermore, during the multiple handling of each of the items, they become mangled, torn, folded, punched full of holes, stamped with endorsements, and scribbled upon with pen and pencil. These bedraggled documents are the ones which must be handled at high speed and from which information must be obtained through automatic means.

Consider the next factor—the need for accuracy. The banking business in particular demands very careful accounting procedures. Errors will be made in an accounting system whether this system is composed entirely of human beings or is entirely mechanized. However, it is essential that these errors be found and corrected before they have a chance to affect the customer.

Therefore, any machine designed for this purpose must place a very strong accent on detection and correction methods.

The fifth factor which was mentioned—system changes must be evolutionary—may come in for a lot of argument from those who can clearly see easier ways of doing things and are impatient when faced with procedures which seem redundant or unnecessary. However, the banking business is in very close contact with its customers. Furthermore, commercial check accounting is so standardized that competitive banks really haven't much to offer other than service. Slight differences in service to a customer can often determine where he does his business. Under these circumstances, no bank will change the customer's habit without careful thought and any change is usually brought about in a very gradual fashion.

It is difficult to generalize on the effect which close dealing with the public has on system design. In the specific system in this paper—that is, a system to handle the accounting in the commercial checking business—a number of things are traceable to the influence of the public's habits. For example, it was not possible, at the outset, at least, to spread the job of statement printing over all the working days during the month, since customers were already used to receiving statements at a given time. It was also considered inadvisable to abbreviate the statements—that is, to give to the customer only the net results of his month's transactions. Instead, it was considered essential to continue the present practice of providing a detailed record of all activity in each account. Also, customers have for years been able to obtain a special statement upon request, and this condition was carried over into the automatic system.

It was not possible to standardize on the width, length, or thickness of the check and deposit slips used by the customers. Furthermore, when it was desirable to preprint certain information on checks and deposit slips (such as the account number) it was impossible to restrict the printing of checks to a few suppliers, since many customers have their own printing arrangements and refuse to change them. This has an important bearing on the tolerances which can be expected in such printing operations and on the complexity of printing equipment which can be allowed.

Further effects were felt in the inability to standardize on service charge schedules. The system had to cope with several schedules on regular commercial accounts and also had to recognize the 10-plan (pay as you go) accounts, as well as the student 10-plan accounts requiring a different charge rate. The customer today also is used to obtaining information on his current balance whenever he so desires and this was carried over. Stop payments and holds had to be handled on short notice.

A last but important aspect of dealing with the public is the impossibility of controlling the volume of activity in a given customer's account. The number of checks the customer cashes can vary widely during different times of the year. This is obviously a point which can hardly be changed by any evolutionary process; therefore, the automatic system must be designed to cope with such fluctuations.

The nature of the filing problem is also worthy of note since it does affect the system. A large per cent of the file is affected during each working day in commercial check accounting. In the Bank of America system, for example, approximately one-quarter of the accounts have entries made to them every day, and, on an overall average, each entry involves about four items. Not only is it essential to have a complete record of this activity for use in service charge calculations, error detection, and printing of statements, but the data must be kept collated or at least must be kept so that the data on one account can be rapidly pulled together, to provide a complete record of that account. This requires a high-volume storage system organized in a manner that gives reasonable access time.

The same conditions apply to the filing of actual checks and deposits; *i.e.*, the physical pieces of paper, with the further requirement that it is necessary to "pay the signatures" on the checks, and this is easiest to do when all the checks of the given account are grouped together.

The ERMA Program

Let us now turn to a brief description of the equipment which was developed by Stanford Research Institute for the Bank of America to handle its commercial check accounting problem. The general area of work handled is shown in Fig. 3. The bookkeeping for several branches is handled by the automatic equipment, and communication with this central location depends largely on messenger service. The first system handles the data for 32,000 active accounts, plus 8,000 clearing items, giving the equivalent of about 12 average-sized branches. At least 36 more units are needed and subsequent models will be made to handle 50,000 accounts.

Out of this program came techniques for high-speed handling of checks of assorted weights, sizes, and conditions; a means of automatically reading information from these checks while moving at high speed; and an automatic data-handling system. It is not the purpose of this paper to describe any of these in detail. However, these techniques are of interest here because they do represent the output of a program faced with the restrictions and demands which were discussed above.

The techniques used in the first engineering model of the equipment for this commercial check processing system employed binary coded numbers printed in magnetic ink for automatic reading of data from the documents. Further development work produced a system capable of reading Arabic numerals printed in magnetic ink without recourse to binary coding schemes. The first equipment produced using this character-recognition approach was employed in the traveler's check accounting operations. However, the development relates very closely to the commercial check bookkeeping being discussed, since the character recognition approach will be used in future models of that equipment. More will be said of this later.

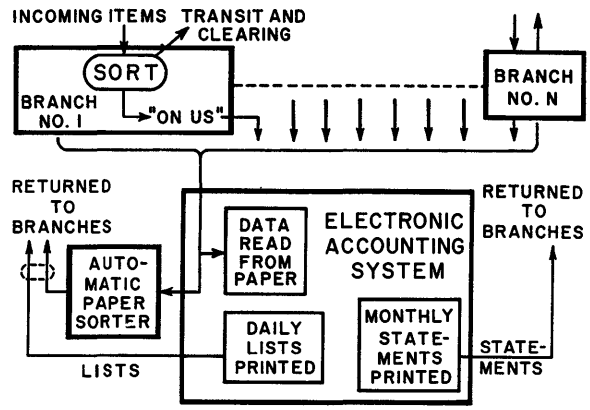


Fig. 3—Central bookkeeping with ERMA.

Let us first consider the ERMA computer. A brief summary of the job being done by this electronic data-processing system is as follows: For each of the 32,000 active checking accounts which it is handling, it records the checks and deposits and keeps the current balance altered accordingly. It scans all incoming items to see if they correspond to "stop payments" or "holds" and to see if they result in overdraft or withdrawal below the hold level. It sorts the incoming items, which are entered in an entirely random order, and collates successive day's work so that all the activity for one account is kept together. It calculates service charges for each account, posts these charges automatically, and prints statements for all accounts. In addition to those operations, which represent the backbone of the work, the system provides lists necessary for analysis and control of the accounts. This includes such things as lists of current balances, lists of overdrafts, stop payments, holds, lists of credits and debits entered during any day, and summaries of the work handled, grouped in order to provide quick diagnosis and correction of errors.

It is a special purpose machine, with a wired program, rather than a stored program. It uses magnetic drum and tape storage and is based on standard vacuum tube circuits and relay switching networks. The system is large, containing approximately 8,000 tubes, 34,000 germanium diodes, and 2,400 relays. Two magnetic drums provide a storage of 3,000,000 bits, and 12 magnetic tapes provide 400,000,000 bits used in a manner which allows storage of 37,000,000 alphanumeric char-

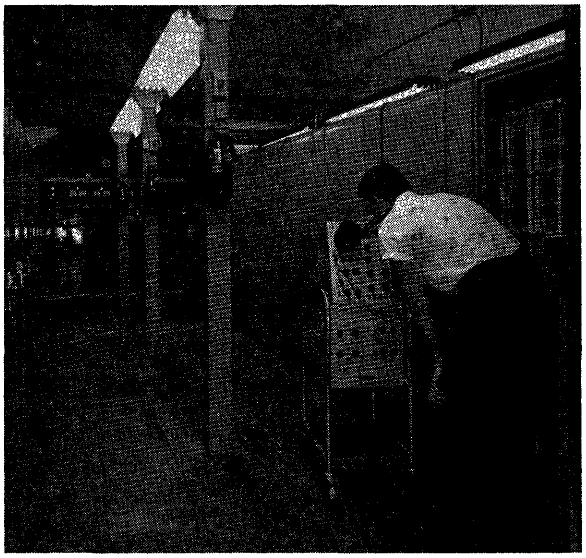


Fig. 4—Main frame.

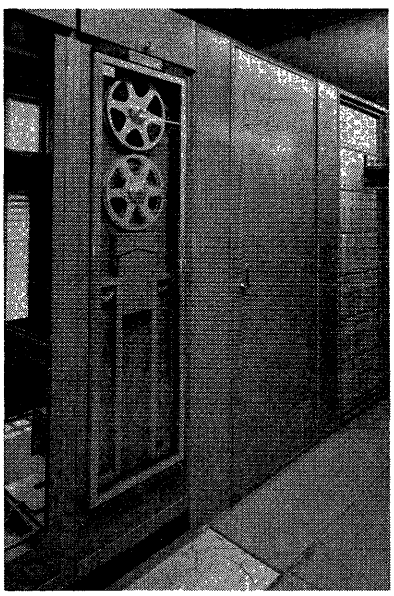


Fig. 5—Tape units.

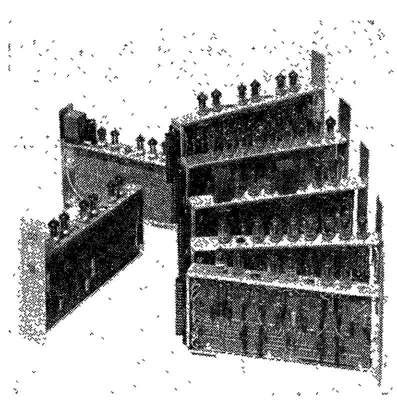


Fig. 6—Circuit packages.

acters. Input is through a combination of keyboards for manual entry of dollar amount and magnetic scanners for automatic entry of account numbers. Auxiliary input for such items as customer name and address is through punched paper tape. The bulk of the output is through a line-at-a-time printer. However, printed outputs for special interrogations can be obtained through the printing mechanism associated with each input keyboard. These keyboards are essentially mechanical adding machines with electrical "read in" and "read out." The system uses excess-3, binary-coded-decimal notation on the magnetic drums and a similar but expanded code on the tapes to handle alphanumeric data.

Fig. 4 gives a view of the main body of the computer but does not show the tapes, inputs, high speed printer, power supplies, or maintenance console. Fig. 5 shows one of the tape units mounted in the tape racks and also shows the front of one of the circuit package racks. Fig. 6 shows several circuit packages. Twenty-four different types of packages are used to cover all logical functions, input/output functions, and communication with memories. Relays are also housed in plug-in packages. Fig. 7 shows the magnetic drums, and Fig. 8 gives a view of one of the manual input stations.

The machine shown in these figures is an engineering model designed with an eye toward production and maintenance. It will go into actual operation and will handle the work of a number of branches. This will allow a thorough shakedown of the operational features of the machine, as well as the engineering aspects, before further models are produced by a manufacturer. The point of view which was adopted in the design of the machine was fairly extreme. It was made to perform nearly all of the functions which could be desired in a system for this purpose. Operational experience will help determine areas where compromises should be made.

The diagram of the work flow through this centralized system (see Fig. 3) indicates that the checks and deposit slips are sorted into account number order after information has been transcribed from them for processing in the computer. Fig. 9 shows the engineering model of the sorter developed for this job. This particular model is a ten-pocket sorter, plus reject pockets, and is used to sort on a decimal basis on one digit at a time. The checks are automatically picked up from a stack at the far end of this picture by a controlled vacuum system. They are then transported between belts at a speed of 150 inches per second. After the checks are picked up and automatically registered against one edge, the binary coded account number printed on each check in magnetic ink is read and interpreted and a memory system is set so that the appropriate pocket gates will be opened when the check arrives. The gating on each pocket is accomplished by controlled vacuum ports in the drums shown above each pocket. The system operates at a speed of ten checks per second. (The techniques have

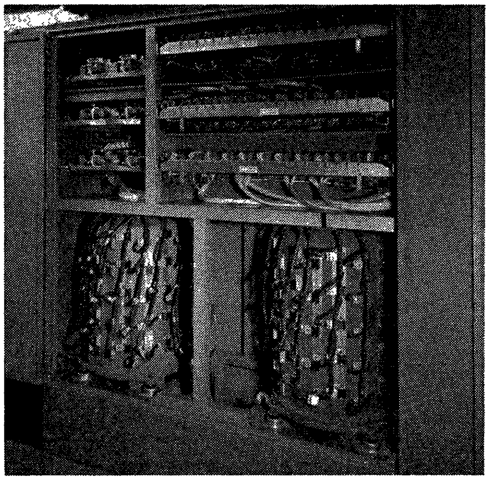


Fig. 7—Magnetic drums.

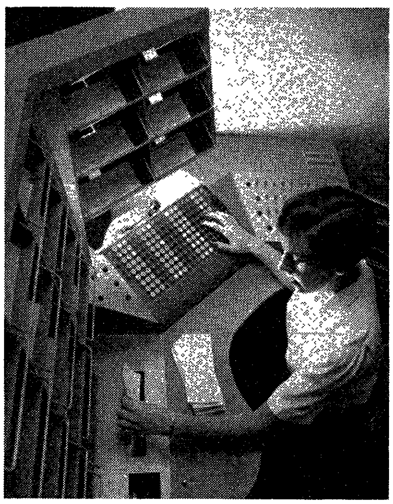


Fig. 8—Manual input station.

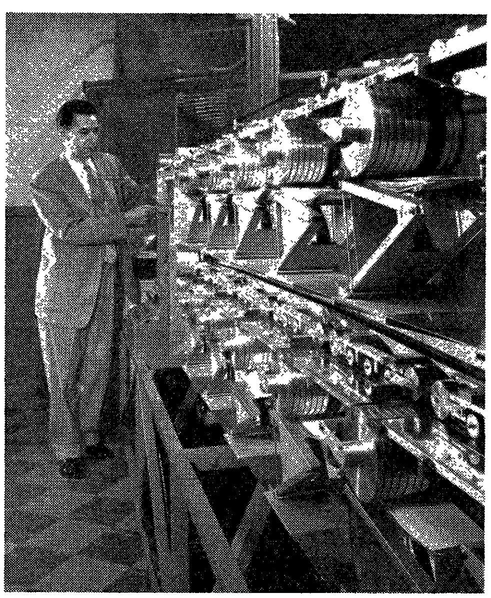


Fig. 9—Check sorter.

recently been extended to handle forty checks per second.) This system works with checks of various sizes and thicknesses; width may vary between $2\frac{1}{2}$ inches and 4 inches, length between $5\frac{1}{4}$ inches and $9\frac{1}{2}$ inches, and a thickness variation between 12 pound paper and punch card stock. Typical checks which are processed by this sorter bear all the earmarks of the public's tender care. They are bent, torn, limp, and disfigured in every conceivable way.

While the work on reading of binary coded numerals was being carried out, plans also progressed for a more difficult but more desirable method, namely, reading numeric characters without the use of binary codes. This is more desirable because it avoids some of the printing problems connected with binary coding, it facilitates visual checking by operating personnel, and it requires less room on the document. Again, magnetic ink was used because it is not affected by over-stamping and other visual mars. Reading is accomplished in the first developmental system with a single magnetic head. Head width is considerably greater than the height of the characters being read, thus easing the registration problems during printing and automatic handling of the paper. The system also tolerates a relatively larger skew angle which may be developed either during the printing or reading operation.

An ideal way of obtaining field tests on this method was found in the travelers check operation, in which a punched card must be produced for every travelers check after it is returned to the bank. This card is used in the present travelers check accounting system. This application was ideal for a test of character reading in the sense that the printing could be carefully controlled through one or two commercial printers, and the characters normally used on these travelers checks are relatively large—about 6 characters per inch.

Fig. 10 shows a view of the travelers check scanning equipment which automatically transports travelers checks past the magnetic reading head, picks off waveshapes characteristic of the numerals being scanned, decodes these numerals by sampling the wave form and converts these samples into binary code. In one scan of the check, the serial number of the check is read, as well as a digit denoting the dollar amount and a digit used for a "nines check," and a punched card is produced. Checks which are not satisfactorily read for any one of a number of reasons are automatically fed into a reject pocket and a reject card is punched and off-set in the stack for easy identification. The electronic portion of this scanner consists of about 170 tubes. Approximately 50 of these tubes are used for the actual decoding operation. The others are used for control, arithmetic checking, and buffer storage for the punch. The small cabinet with a large number of control knobs in Fig. 10 is a special test device used to detect errors during a trial run in the laboratory. The error rate was so small that it could not be detected adequately without special gear.

Fig. 10—Travelers check scanning equipment.

This unit is finished and will be put into operation as soon as checks bearing the magnetic numerals are in use in sufficient numbers. Preliminary tests using a specially printed run of 100,000 travelers checks gave the following results: Out of the first 10,000 check passes, two errors were found; *i.e.*, cases in which the scanner thought it produced correct results but which were proven incorrect by the external detection equipment. Additional error detection circuits were added to the decoder and no further errors occurred during the next 90,000 check passes. Rejected items, *i.e.*, cases in which the scanner said that it could not read properly, ran as follows: Two per cent rejects occurred during the first 38,000 checks. (Analysis showed that 1.4 per cent were rejected because of malformed experimental printing type.) During the remaining check passes, 0.76 per cent rejects occurred, excluding the cases in which bad type caused reject. Currently, in the daily operation of the scanner, the machine generally rejects about 0.1 per cent of the checks and seldom rejects as much as 0.2 per cent.

One further area of development should be mentioned. Automatic paper handling systems of much wider application can be built if the dollar amount can be printed on the check in a manner which can be transcribed automatically. These dollar amounts might be written on the check by check-writing machines in large corporations, or they might be post printed by the banks during the first manual processing after the items return to the banking system. In any case, this poses the the requirement of making the printing mechanism relatively inexpensive, hence printing tolerances cannot be held very closely and the character recognition system must be able to cope with these wider tolerances.

This post printing and reading, which has great potential in such areas as the department store and oil company credit business as well as in check handling, is now in the development stage.

Conclusion

This paper has attempted to show the magnitude of the check handling problem in national banking and the factors which influence the design of a system for automatically handling this problem. It has implied but not attempted to prove that bank check handling problems are fairly representative of financial accounting in general, particularly credit handling and centralized money management. It has implied, and it should again be emphasized here, that means for automatic handling and transcription of basic documents are an essential part of any system designed for this kind of operation. The paper has described an experimental program which resulted in equipment for actual use. This equipment by no means represents the ultimate, but it does represent progress and a good starting point for future development in the area defined by the general principles of financial accounting.

Acknowledgment

The authors wish to emphasize that the program discussed in this paper represents the work of many other people. This paper is intended to be introductory in nature and will be followed by more detailed papers written by members of the staff, covering the technical aspects of the program.

Discussion

J. D. Mountain (Mountain Systems, Inc.): What kind of numbers were printed on the checks in magnetic ink, the account numbers or dollar amounts?

Mr. Noe: The first system has just the account numbers printed and these are in a coded form. Subsequent systems will also include dollar amounts and hence replace most of the manual input operations. On travelers checks the job is simplified, there being only four dollar amounts to deal with; number one for ten, two for twenty, three for fifty, and four for a hundred.

J. Svigals (IBM): Does ERMA satisfy the need for economy? What is its cost and how much will it save?

Mr. Noe: The Bank of America people have not released actual figures on costs and savings. However, they would not have gone through with the program had they not foreseen adequate returns on the investment.

T. C. Morrill (Liberty Mutual Insurance Co.): Can ERMA be applied to other types of work than commercial banking and, if so, is it available for sale?

Mr. Noe: The equipment is not now available for sale. The question of future sales will be handled jointly by the bank and the manufacturer who produces it. The system, as it stands today, is very special purpose in that it handles only this particular dollar accounting job; but we who have been working with it feel it can be extended without too much difficulty to handle other jobs in the dollar accounting area.

N. J. Dean (Laboratory for Electronics): Are all checks posted as individual entries or are they posted as a single list total when there is a large enough number on one account?

Mr. Noe: They are posted as individual entries. For one thing, most of the items come in unsorted; in other words, in entirely random order. Also they do need posting individually at this stage because it is necessary at the input of the machine to protect it against human errors. For this reason the operator lists each item and accumulates the dollar amount, comparing the total with the total on the list which came with the batch. This insures that she didn't punch the wrong key or that an item didn't fall down the elevator shaft while being carried to the central ERMA office.

The Manual Use of Automatic Records

A. G. OETTINGER†

Summary—**Records that are to be processed automatically necessarily must be stored in media accessible to a machine and in codes suitable for machine interpretation. Consequently, such records may be not readily accessible to the unaided human clerk. Steps that may be taken to provide adequate facilities for the occasional manual interrogation and processing of automatic records are suggested and examined.**

In THE BROADEST sense, access to any automatic system is manual, human intervention being required in the operation of every system, if only to push a button to start it the first time. The problem of concern here thus arises from questions of degree. An automatic system operating unattended has its own tempo, a tempo often several orders of magnitude greater than the fastest human tempo. Information can be said to be processed automatically, if processed at the system's own tempo, and manually, if at a clerk's tempo. In general, the performance of any fully specified process can be controlled by circuit elements or by program steps, and hence may be regarded as automatic. It is chiefly the unforeseen, the rare, or the random which necessitate some form of manual intervention in the automatic process. A given system might thus be operated automatically at one time, manually at another. For example, the printing of all the information in the storage unit of a system by a printer operating at the system's own tempo, is an "automatic" process. The same information, displayed item-by-item on a monitor typewriter or some other display device, at the tempo of a button-pushing operator can be said to have been obtained manually. Between these two extreme cases, there is a whole spectrum of information processing rates and of system efficiency.

The automatic tempo of a system and the influence of manual intervention on the operation of the system depend on several factors; those of greatest importance here are the rates at which information can be introduced into the system, processed, and extracted, and the size of the information files. The files will be assumed to be large and the rates to be determined chiefly by the characteristics of the input, output, and storage facilities of the system. The problem of manual access to automatic records is therefore but one facet of the broader problem of designing data processing systems with optimal input, output, and storage facilities. It will be useful to analyze some of the characteristics of this broader problem in order to appreciate the particular effects of manual intervention on the efficiency of a system.

Storage devices are commonly classified as "random-access" or "serial-access." This classification unfortunately is confusing, for it is not precise, and it masks some important and useful distinctions. Storage devices now lumped as "random-access devices" may meet any one or more of the following essentially distinct and independent criteria:[1]

1) Information can be extracted from the device at random as rapidly on the average as if extracted in the order natural to the device.
2) The average access rate to information in the device is negligible with respect to the rate at which the information is required for use by associated equipment.
3) The basic unit of data (character, word, or block) is located by specifying suitable coordinates rather than by search.

The first criterion is based on a characteristic pertaining exclusively to the storage device. For any storage device, the access time to an item depends on the location of the item itself and often also on the location of the previously selected item. There exists, consequently, a preferred or "serial" order of selecting items, for which the average access time per item is a minimum. A device whose average random access time is comparable to its serial access time, is now characterized as "random-access."

As for the second criterion, if any arbitrary item of information can be supplied from storage as needed by the system with a negligible delay, the storage device again is termed "random-access" regardless of its specific physical nature. The third and last criterion is independent of time and often of the physical nature of the storage device as well. Again, however, it is common to designate any storage device addressed by coordinates as a "random-access" device.

Iverson has proposed to clarify the terminology by restricting the designation "random-access" to devices satisfying the first criterion. He adopts the following more precise terminology to characterize storage devices:

random-access or *serial-access*, according to whether or not a storage device satisfies the first criterion;
immediate-access or *delayed-access*, with respect to the second criterion;
coordinate-access or *search-access*, with respect to the third criterion.

The distinction between immediate access and delayed access can be applied not only to the storage elements of a data processing system, but to the whole system as well. In this case, the characteristics of the

† Instructor in Applied Mathematics, Computation Laboratory, Harvard University, Cambridge, Mass.

[1] The author is indebted to Professor K. E. Iverson for suggesting these distinctions.

system input and output devices must also be taken into consideration. For example, consider a system provided with immediate-access internal storage and single input and output channels. Since the internal storage is immediate-access, the over-all access rate is limited effectively only by the input-output channel. Suppose that the delay between the receipt of an inquiry at the input and the delivery of an answer by the output[2] is well below the patience break-point of the inquirer. Then, provided that the interval between successive inquiries is greater than the pipeline time, the system may be termed an "immediate-access" system. On the other hand, if inquiries arrive at a higher rate, one may arrive while the single channel is still tied up by the previous inquiries. The system, in this case, is a delayed-access system. Furthermore, if this input channel is normally preempted by routine operations, any inquiry must interrupt the routine operation, delay its completion and, because of differences in tempo, greatly lower the efficiency of the system.

It is convenient for further discussion to define a measure of the efficiency of an automatic system as a data processing tool, in terms of its advantages in speed, cost, accuracy, etc., relative to alternative tools. Such a measure might well be defined by analogy to the definition of the mechanical advantage of rope and pulley fame.

Inasmuch as new automatic systems should be improvements on existing manual or automatic systems, it seems natural to measure the automatic advantage of new systems relative to the existing system. However, any other suitable reference systems may be used. For example, if operating rate were the only significant factor, the automatic advantage (AA) of a new system might be defined as

$$AA = \frac{r_n}{r_r}$$

where r_n is the operating rate of the new system, and r_r that of the reference system.

Since, in reality, operating rate is but one of many factors of importance in the evaluation of a system, the automatic advantage might be better described as

$$AA = \frac{f(r_n, c_n, l_n, p_n, \cdots)}{f(r_r, c_r, l_r, p_r, \cdots)} + I,$$

provided that the form of the "value" function f can be specified. The variables r, c, and l represent respectively the operating rate, cost, and load factor of the system, and p the cost of system personnel. Many other variables could and should be included in a realistic expression. The additive term I is a handy adjustable factor representing intangible benefits, and may be used to set the automatic advantage at any desired value.

In the absence of an explicit definition of the form of the function f, the automatic advantage is barely more than a qualitative indicator of efficiency. However, in considering the effects of manual intervention on the efficiency of an automatic system, the resulting change in efficiency is of prime interest. A knowledge of the exact magnitude of the automatic advantage is therefore not essential if this measure is to be used only as an aid to clarity in exposition. Parenthetically, it should be clear that the explicit specification of the form of the function f is equivalent to a precise knowledge of the relative worth of two systems. That a problem of such importance has not yet been solved to anyone's complete satisfaction is an indication of its complexity.

On the basis of the more refined definitions of access characteristics and the measure of efficiency presented in the preceding paragraphs, it becomes possible to consider in some detail the effects of manual intervention on the efficiency of automatic systems. For the sake of clarity, a specific example will be used. This example will be drawn from banking applications.

Banks today are keenly interested in developing automatic systems for checking account bookkeeping. Reduced to its simplest terms, the problem is that of keeping a record of each depositor's balance, increasing it when deposits are made and reducing it as checks are drawn. The problem is complicated by the need to handle and store the checks themselves, from the time they arrive at the bank until they are returned to the depositor.

In the course of a day's work, the bookkeeping department of a bank branch handling some 30,000 accounts of different types may post 65,000 debits and 7,500 credits. During business hours, the same department may also be required to answer the inquiries classified in Tables I and II, as well as to perform the special procedures listed in Table III. The figures in these tables are averages based on actual operating statistics.

It is striking that no more than one per cent of the accounts are affected by the inquiries or special procedures. The volume of special cases is dwarfed also by the volume of routine debits and credits by a factor of two hundred. In the initial confusion attending any large scale systems study, it is natural to concentrate attention on the major production operations. Distracting special cases tend to be neglected until a preliminary study has led to a sketch of a system meeting the major operating requirements.

A preliminary study of this kind might well lead to the system sketched in Fig. 1. The basic file maintenance unit is assumed to be a general purpose computer part of whose time is to be allocated to the check-bookkeeping operation. The account records are assumed to be kept on magnetic tapes, and a single tape input mechanism ① is provided to make these records accessible to the file maintenance unit. The natural order of access to records stored on a given tape is

[2] This delay has been called the "pipeline time" of the system in "File Reference," by J. A. Postley, Rept. P-691, Rand Corp., 1955.

TABLE I

Daily Inquiries Classified by Source

Teller	47
Customer by telephone	49
Others (credit dept., loan dept., etc.)	65
Total	161

TABLE II

Daily Inquiries Classified by Type

Does a given person have an account?	5
What is the balance of account X?	85
Has a certain check been posted?	7
Has a certain deposit been posted?	5
Is there a stop-payment order on account X?	13
Is there a hold order on account X?	34
Others	12
Total	161

TABLE III

Special Transactions (Daily)

New accounts opened	39
Changes of address	12
Accounts closed out	27
Statement requests	22
Overdrafts	58
Stop-payments initiated	36
Stop-payments revoked	18
Holds initiated	56
Total	268

therefore the order in which they are recorded on this tape. However, any one of the set of tapes may be placed on the input mechanism as readily as any other. Debits and credits ordered in the natural account order are presented at the input ③. Physically, the debit and credit file might be checks and deposit tickets imprinted with a machine-readable code, or a magnetic tape on which the proper entries have been recorded previously. The output ② is an updated magnetic tape account file.

The system, as sketched, provides the file maintenance unit with immediate access to account records. In part, this is insured by the common ordering of the two input files. Furthermore, between thirty and fifty per cent of all accounts are activated in the posting operation. Tape travel time between active accounts can therefore be considered negligible, especially if provision is made to advance both input files to the next account while one is being processed. The system, therefore, is of balanced and efficient design, and has an automatic advantage determined chiefly by the cost of the computer time used for the check-bookkeeping operation, the cost of the key-punching and sorting required to prepare the debit and credit file for processing, and the gains in speed and accuracy provided by automatic operation.

It will be assumed now that the automatic advantage of this system is attractive so far as the performance of routine posting operations is concerned. It remains to consider the effects of the inquiries and special procedures listed in Tables I, II, and III. A number of these cases may be treated very simply by adding to the basic system the elements indicated by broken lines in Fig. 1. An additional input mechanism ④ is provided to introduce inquiries and special transactions into the processing unit from a file ordered, again, in the natural order of the account file. Magnetic tape could be used for this "special" file, but since few accounts are affected, paper tape or cards might be adequate. A "service file" ⑤ is provided at the output to deliver answers to inquiries and to bring special transactions to the attention of the operating staff.

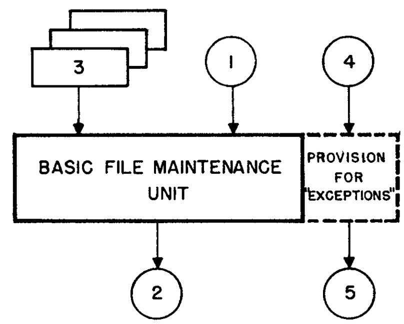


Fig. 1—File maintenance and consultation system.
Basic System: 1) Ordered account file—input. 2) Ordered account file—revised. 3) Ordered transaction file.
Provisions for "Exceptions": 4) Ordered inquiries and special transactions. 5) Ordered service file.

The provision of such additional facilities increases the cost of the system, but as operating speed is virtually unaffected, and the system is made more powerful, the automatic advantage probably suffers but little. In principle, this procedure is applicable to any transaction whose consummation may be delayed until the next regularly scheduled pass of the proper account tape. In practice, the frequency of these transactions must be sufficient to warrant the provision of the added equipment. Among the special transactions listed in Table III, the opening and closing of accounts, and address changes probably satisfy these conditions. Requests for statements are clearly a borderline case. A one day delay can usually be tolerated, but if a large depositor has been favored with a surprise visit by his auditor, his request for immediate delivery of a statement cannot lightly be deferred. Overdrafts must be reported in time for management consideration and action prior to the daily deadline for the return of checks drawn against insufficient funds. Stop-payment and hold orders[3] must be noted at once. A request by a customer or a department of the bank for the balance of an account cannot be deferred for a day. The additional facilities indicated in Fig. 1 are inadequate to these contingencies, and other solutions must be sought.

[3] A stop-payment order issued by the maker of a check instructs the bank not to pay the check when it is presented. A hold order, issued by a teller or a bank officer freezes all or a portion of the balance of an account. Tellers issue hold orders when cashing or certifying a check, to maintain a balance sufficient to cover the check when it reaches the bookkeeping department.

Tellers and customers frequently telephone the bookkeeping department to ask for the balance of certain accounts. As the figures of Table II indicate, such calls are likely to occur eighty-five times a day, and they represent fifty per cent of all inquiries. If the system of Fig. 1 happens to be processing the very reel of tape on which the account in question is stored, an answer conceivably may be obtained in a matter of seconds. In general, however, the system will be processing some other reel, or working on an entirely different problem.

A tape file containing all the account information for 30,000 depositors now customarily held on ledger cards is likely to require from ten to fifteen reels of tape. To answer an inquiry using the system of Fig. 1, it is necessary manually to select the proper reel, interrupt the operation of the system, remove the reel being processed, mount the new reel, set the machine to the proper search procedure, and start it on its way. Surely it is conservative to assume that at least two minutes will be required for these operations. Conceivably, the inquirer will have his answer within three minutes. This delay approximates that presently required for the manual consultation of a file of ledger cards, and therefore seems acceptable. Thus, for reasonably spaced inquiries, the system is of immediate access. Eighty-five inquiries, at two minutes each, do however require three hours a day of system time. In comparison, the time required to perform the 70,000 routine daily postings may be approximately two hours. Answering a relatively minute number of inquiries thus adds some 150 per cent to the estimated operating time. If, as assumed, it had been planned to parcel the time of the file maintenance unit among several applications, the number of applications must be reduced correspondingly. The result is an increased allocation of operating costs to check-bookkeeping, with the consequent reduction of the automatic advantage of the system in this application. Processing the inquiries in this way offers little, if any, compensatory advantage over the purely manual consultation of ledger cards. Equivalent manual labor is required, and the delay in answering may even be increased. Indeed, it seems to be true in general that *file consultation* alone does not warrant the use of automatic equipment. In the system sketched in Fig. 1, as in many systems, the automatic advantage in routine posting or *file maintenance* must be substantial enough to leave a net automatic advantage when file consultation is taken into account.

A possible variant of the system of Fig. 1 is one with several tape input mechanisms. Indeed, it is more common for machines to have several tape mechanisms than not. Selection among tapes can then be automatic, and relatively rapid coordinate access to any of these tapes is possible. To the extent that manual tape handling is eliminated, the provision of multiple tape mechanisms tends to increase the automatic advantage of a system. However, since the automatic advantage is a function also of equipment cost, a net reduction of automatic advantage may result from an increase in the number of tape mechanisms. The most efficient number of tape mechanisms depends on too many factors to be specifiable here, but systems with multiple tape mechanisms are probably satisfactory in many applications. A corollary improvement can be obtained by separating the file consultation and file maintenance operations. The provision of an independent file search unit can relieve the file maintenance unit from most interruptions. A conflict occurs when both units require the same tape at once, but interruptions due to this event are far less frequent. For large files, access time to records in a multiple-tape system probably remains comparable to manual access time to ledger card files. Furthermore, the provision of a file search unit adds significantly to the cost and again may reduce the overall automatic advantage.

Another method of handling inquiries is based on the periodic printing of selected contents of the automatic file to provide a readable file for manual use. The information in a given printed record cannot include changes in the automatic file made after printing. Nevertheless, if the file is inactive, or if approximately current information is adequate, this limitation may not be serious. In principle printed files could be produced as frequently as might be necessary to provide information of adequate currency. Printing time and cost do, however, limit the practicable printing frequency. With a large file, printing time and cost may be economized by setting a long interval between successive printings of the whole file, while providing more frequent printings of the information in active accounts only. These "incremental" prints then serve to keep the main printed file current. Unfortunately, consultation of the printed file becomes ever more difficult as incremental prints are added to the original. Especially where file activity is high, this may lead to a manual file maintenance problem comparable to the one which the automatic system was intended to solve. Thus, in general, file printing is probably more of an expedient than a satisfactory solution.

The examination of alternative methods of treating inquiries and special transactions presented in the preceding paragraphs has been based on the tacit assumption that the system sketched in Fig. 1 is fundamentally sound. However, it is clear that the automatic advantage of the system in file consultation is not impressive. The system as a whole has a net automatic advantage only if it excels in file maintenance. It remains to examine alternative basic designs.

One important alternative is the use of random-access storage devices instead of tapes. Another is the use of a new type of unit records. A third is the operation of several smaller, slower processing units instead of the one large, fast one. Innumerable methods based on varying degrees of integration of clerks and machines suggest themselves. Only the first two alternatives will be treated in any detail here.

Random-access storage devices of the drum or of the "juke-box" type offer little advantage over tapes in the routine file maintenance operation, so long as the ordering of transactions is of negligible cost. Serial-access times to tape, drum, or "juke-box" being comparable and immediate in this case, the lower cost of tape gives it a decided edge. If, on the other hand, the use of sorting to arrange incoming transactions into the natural order of a serial-access file is precluded by cost, only random-access devices remain of immediate-access and their relative automatic advantage is enhanced. A detailed analysis of the influence of sorting requirements on the relative automatic advantages of serial- and random-access storage systems is beyond the scope of this paper, but the interested reader is referred to the works of Hosken and of Gregory elsewhere in this volume, and to those of Ashenhurst and Seward.[4]

In the treatment of inquiries and special transactions the random-access devices potentially offer substantial advantages. Insofar as they are used with coordinate access, any addressable unit of information is very nearly as accessible as any other. Hence, the interruptions of file maintenance activities for answering inquiries can be several orders of magnitude shorter than with tape. The automatic advantage of a system with random-access storage, regarded as a function of operating rate only, can thus be significantly greater than that of a tape system.

Despite these advantages, random-access storage is not a panacea. Interruptions of file maintenance activities, although shorter, are still present and can be eliminated only by the provision of multiplexing facilities. Other things being equal, random-access storage devices have little if any advantage over serial access devices when search access rather than coordinate access is used. Finally, the cost of random access storage facilities for the estimated 2×10^8 binary digits in 30,000 accounts far exceeds that of ten to fifteen reels of tape. Thus, if regarded as a function of operating rate and storage cost only, the automatic advantage of a system with random-access storage may be appreciably smaller than that of a system with tape. The inclusion of sorting costs in the evaluation of the automatic advantage favors random-access devices, but serial-access devices are by no means excluded from consideration.

There is a real possibility that random-access devices can play an important role as auxiliary immediate-access files in systems using tapes to store the main files. Thus, for example, information regarding holds and stop-payments might be stored in such a device, where it would be readily accessible both to tellers and other inquirers and to the file maintenance unit. Since only ten per cent of the accounts are affected by holds and stop-payments at any time, this possibility is economically attractive. There is reason to believe that inquiries tend to affect a relatively small proportion of the accounts, and that generally the same accounts are affected. Should this be true over sufficient periods of time, there might be some advantage in segregating these accounts into a special file. This file might then be held in random-access devices, while the bulk of the account file remains on tape.

It has been suggested that banks do not really need to record all transactions in conventional ledgers, that it is sufficient to maintain immediate-access records of balances and to print transaction details only at statement time. In this case, the immediate-access storage requirements would be reduced by a factor of ten or more. The automatic advantages of both serial- and random-access storage devices are apparently increased thereby, although in different proportions. This apparent increase is due, at least in part, to an improper choice of the reference system used in determining the automatic advantage.

Many banks actually have adopted the suggestion to maintain immediate-access records for balances only, in the form of the so-called "post-to-check" plan. In brief, this plan eliminates the ledger card file altogether by recording the current balance on the face of the last check or deposit ticket posted. The checks themselves thus provide an immediate-access record of current balances, as well as a delayed-access[5] record to transaction details. In some cases, these details are transcribed to statements before they are issued, in others the statement merely gives the balance. For many reasons, the application of this plan so far is largely confined to "special checking" accounts. However, in evaluating the automatic advantage of an automatic system in which only balances are to be stored with immediate access, it is clear that the reference system used in the evaluation must be one akin to a post-to-check system, and not one using an immediate-access file of complete transaction records. It makes little sense to simplify a procedure merely to obtain an apparent increase in the automatic advantage of a proposed automatic system, when the same simplification of an existing system might yield an equal or greater automatic advantage.

A second basic design alternative to that of Fig. 1 would depend on the use of unit account records, rather than of continuous media of the tape or drum type. The major unit records now commonly used in automatic systems are the well-known punched cards. These cards have been superseded by magnetic devices in some applications chiefly because of their low information density. However, low information density is a characteristic of accepted punched card technology, rather than an inherent property of unit records.

[4] See Hosken, J. C., "Evaluation of sorting methods," Gregory, R. H., "Document processing." Ashenhurst, R. L., "Sorting and arranging," *Theory of Switching*, Progress Report No. BL-7, Harvard Computation Lab., 1954; Seward, H. H., "Information sorting in the application of electronic digital computers to business operations," Report R-232, Project Whirlwind, M.I.T.; 1954.

[5] With respect to inquiries.

A card of the size of a standard punched card could easily hold 20,000 binary digits on one magnetically-coated side, and a $9\frac{1}{2}'' \times 9\frac{1}{2}''$ ledger card commonly used in checking account bookkeeping could easily hold 80,000 binary digits. Storage capacity of this magnitude vastly exceeds the requirements for any one account. Furthermore, the uncoated side of the card could readily be used for essential reference information in ordinary printed form. When file maintenance operations can readily be performed with the use of ordered files, moderate card feed speeds are adequate to provide immediate access to a magnetic card file. If the technological problems of feeding the cards and of reading the magnetically-recorded information can be solved economically, a system based on these cards may well rival the automatic advantage of tape or drum systems in ordered file maintenance operations.

File consultation in a magnetic card system could be performed by methods very similar to those now in use with ordinary ledger cards. Inquiries now answered by clerks within three minutes would continue to be answered at the same rate and at the same low cost. Objections regarding accidental disarrangement of the files might well be overcome if as much effort were devoted to the improvement of unit record files as has been spent on the development of tape and drum devices. For example, the use of a device akin to the familiar rod threading cards in a library catalog could prevent the complete removal or the accidental disordering of cards, while still permitting ready consultation. Urgent but infrequent transactions such as the noting of stop-payment or hold orders could be performed by temporary manual entries on the face of the cards. A permanent record can be made later by introducing appropriate entries into the transaction file for normal processing by the file maintenance unit. Such a combination of the best features of manual and automatic operations seems worthy of careful investigation.

Many other alternatives will readily come to the mind of the reader and, all in probability, many of the methods described here are under consideration in scientific or industrial laboratories. However, there are no published experimental data or records of experience which, to the author's knowledge, could justify concluding that the problem of manual access to automatic records has been solved.

In conclusion, it should be noted that many of the design problems examined in this paper have their source, at least in part, in inadequate specifications of the operating requirements of data processing systems. Accurate information regarding the source, distribution, and frequency of inquiries is notably lacking. No reliable specifications are available regarding exactly how promptly certain inquiries must be answered, or regarding just how dire the dire effects of delay would be. The marginal cost of treating exceptions is frequently not considered. For example, the cost of permitting an occasional overdraft or of reimbursing a depositor for the amount of a check paid in spite of a stop-payment order should be weighed against the cost of guaranteeing against occasional lapses. It may then be decided that protection against fraud is worth any cost, but the value of this protection must then be included in the determination of the automatic advantage of a proposed system. A lack of reliable data in cases where the distinction between milliseconds and minutes or hours is a matter of hundreds of thousands of dollars, leaves many a crucial decision to the whim of hunch or pressure.

ACKNOWLEDGMENT

The author is indebted to J. Lewis Nungesser and Joseph Perret of the Philadelphia National Bank, to his colleagues at the Computation Laboratory, and to many others for stimulating discussions of invaluable assistance in the preparation of this paper.

Discussion

N. Dean (Laboratory for Electronics): What percentage or random inquiries refer to detail items rather than balance summary information?

Dr. Oettinger: I have a feeling that probably a majority—well, certainly more than 50 per cent—refer to balances, and I think probably a most substantial part refer to balances only.

N. Dean (Laboratory for Electronics): What percentage of special transactions must be handled immediately; that is, before end of sorted input?

Dr. Oettinger: I think that, in general, transactions like holds require immediate action but again reference is to the balance only. As far as the delay allowed for some of the others, this is one of the questions to which a detailed answer would be appreciated by myself as well as you.

D. R. Klusman (American Tel. & Tel. Co.): Are you aware of any large scale storage systems under development that appear to meet most of the requirements satisfactorily, including random access of a reasonable rate?

Dr. Oettinger: I am not the expert in this field and perhaps Mr. Ridenour's paper this afternoon or tomorrow might give some of the clues.

B. W. Taunton (First National Bank of Boston): Will you ask the members of this audience how many of them actually check the detail items listed in their monthly statements furnished them by the banks with respect to their own checking accounts?

Dr. Oettinger: Raise your hands; everybody who looks at their detailed items. (Vast majority of hands raised.) I don't know whether this is a reflection on the trust in which the banking industry is held, but perhaps it is because you work on the inside of automatic machinery.

G. W. Patterson (University of Pennsylvania): You stated that only one per cent of the accounts in a bank are involved in special inquiries. Is this on a daily basis? How does this percentage increase as longer intervals of time are specified?

Dr. Oettinger: This is a very good question and a very important one. The figure of one per cent is a daily figure and there are reasons to suspect—this is only a hunch—that in certain types of inquiries the same accounts are involved day after day. It would be nice to have a definite answer. If the same accounts were involved they could be separated in a special file. This is one of the design parameters for which I would like to know a definite value. I just don't know the answer to your question but I hope an answer will be forthcoming.

C. Kagan (Western Electric Co.): Would you say that economical maintenance of records in several sequences compels the use of random access devices; assuming a large file of 100,000—80 digital records, 100 transactions a day, that may have to be posted in several of the file units within a particular sequence as well as in the same number of file units in all other sequences?

Dr. Oettinger: I have a rough idea of what that means, but the only answer I could give would be as involved as the question was. I would be happy to see Mr. Kagan privately. The answer would require a study of multiple ordering, and the relationship of this problem to the whole storage device problem is extremely complicated. The only thing I can say is that there are no good answers to this, and that anyone who claims to have an answer is either deluded or a good salesman.

J. Smith (Smith Mfg. Co.): What are examples of coordinate and random access machines?

Dr. Oettinger: Well, a random access device—for example a magnetic drum—is a random access device because the information can be obtained at random from any position on the drum in roughly the same time as if in natural order. Coordinate access refers to the direct selection of individual items without search. Let's say we are talking of words; with coordinate access, every word can be selected independently from all the others. Many of the high speed drums are accessible in this form. On the other hand, there are others where only blocks are accessible by coordinates and words within the blocks are accessible only by searching the blocks. These characteristics are independent of one another; one could use a random access device with either coordinate or search access.

J. Smith (Smith Mfg. Co.): If the memory core matrix is coordinate access, what is used to indicate which coordinates to use?

Dr. Oettinger: The intersection of a set of wires specifies a core; this is actually one of the most graphic types of coordinates.

H. Groelinger (National Bureau of Standards): You seem to imply tape searching usually is a function of special inquiries. These, however, never decrease but steadily and consistently increase. If this is coupled with a business' desire to increase its service to its clients, would you not conclude that the adverse effect of "tape searching" can be ignored?

Dr. Oettinger: If the gentleman wants to see me I would be glad to debate this issue.

Evaluation of Sorting Methods

J. C. HOSKEN†

Introduction

I HAVE BEEN asked to report to you on the present state of the art of sorting. In order to do so, I have obtained help and information from more people than I can thank individually. My former employers, Arthur D. Little, Inc. have allowed me to spend many hours of their valuable time, and my present employers, The Farrington Manufacturing Co., of their, I am glad to say, even more valuable time, on sorting out the information received.

Definition

For my purpose, sorting means putting things in order according to some numerical, alphabetic, or geographical key. The things sorted may be mail or checks, which in themselves should arrive in one piece at a specified destination, or items of information which may be transferred from one medium to another without loss.

Importance

Sorting is an important subject. To lower costs per unit of output, people usually increase the size of their operations. But under these conditions, the unit cost of sorting, instead of falling, rises. Sorting is a major reason for using punched cards. Up till now it has often been a major reason for not using magnetic tapes and otherwise high speed systems.

† Farrington Mfg. Co., Industrial Center, Needham Hgts. 94, Mass.

Why Sort?

The usual reasons for sorting are: to enable information to be found easily, as in a telephone directory; as a step in bringing together pieces of information, for example, in reconciling incoming checks with billing information; to bring files up to date, as in an inventory system; in sorting mail and handling checks; in making a summary report.

How Sort?

In order to get sorting done expeditiously, you can either hire someone who knows the alphabet, or, as they do in the post office, someone who knows geography and is a good pitcher; you can use filing devices, like card indexes; you can build a specialized machine like Erma's sorting mechanism [1]; you can use punched card sorters; you can use electronic clerical machine sorters.

Techniques

In all the various sorting systems, a relatively limited number of techniques is used. These I propose to describe in turn.

Terms

Before discussing the various techniques, let me define one or two terms for my purpose.

An *item*, sometimes called a *record*, is a unit which is being sorted. It may be a letter in the mail, a check or the details of a transaction, such as what Mrs. Murphy bought at Filene's yesterday, or the number of parts #3792 extracted for job #767.

The *key* (or *designator*) is the information on the item which indicates the order in which it is to be sorted; *e.g.*, the address on an envelope in the mail, the part number in an inventory system, the catalog number in a mail order transaction, the subscriber's name in a telephone directory, or the town or district in a fulfillment system.

A *string* is defined as a number of items whose keys are in ascending or descending order.

A *block* is a number of items read into a computer in one operation in some systems.

Distribution

The simplest technique and probably the oldest is distribution. This is the system used in the post office where mail is sorted into geographical categories. The number of categories in each sort is limited by the accuracy of the mail sorter's aim, and is usually about twenty. Each of these categories is subsequently broken down into 20 more and so on.

Digital

The next technique is generally known as digital, but also as column, character, decimal, alphabetic, or diverging sorting. This is the method normally used in sorting punched cards.

Merge or Mesh

The next technique is known as merge, mesh, or collate. This is the most commonly used in machines using magnetic tape.

Inserting, Address Calculation, Selection, Exchange

Several other techniques are practical in dealing with information carried in a relatively high-speed memory, such as a magnetic drum. One of these is the method of inserting, a special form of which is called address calculation; another is the method of selection, in one form called "nth" degree selection; another is the method of exchange.

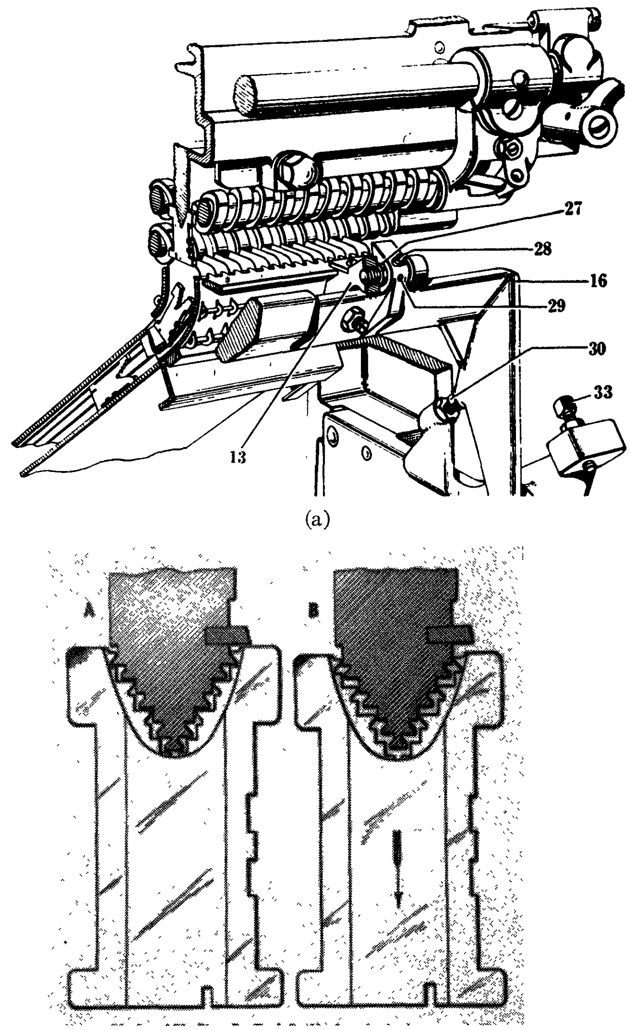
Avoid Sorting

The last important technique is to avoid sorting altogether. This can be done by making somebody else do it. An example in ordering parts from inventory is to supply a form containing all the part numbers in order and to make the person ordering fill in requirements on this form. Another method is to use a large rapid or random access memory. Several of these are gradually becoming available.

Illustration of Techniques

Distribution Systems

As mentioned earlier, the oldest of the distribution systems is mail sorting. The Canadian post office [2] has recently been bringing this up to date by translating



(a) Intertype matrix bar sorter mechanism.
(b) Intertype matrices showing binary keys.

Fig. 1

the address on the envelope into a coded key which is printed on the back of the envelope in dots and can subsequently be read by photo-electric devices controlling the sorting operation. Other mail sorting systems are being developed in Europe. Another system of this kind which has been at work for a long time is the matrix bar sorter which returns the matrices to the magazine in an Intertype [3] or Linotype [4] typesetting machine [Fig. 1(a) and (b)]. In this case, a binary key is permanently cut in the metal of the matrix and the matrix is slid along the matrix bar till it reaches the first point where the voids in the matrix bar match with the teeth in the key. Here it falls off into a pocket.

Jack Rabinow [5] some time ago developed an ingenious conveyor belt sorter [6] of this general type for the Bureau of the Census (Fig. 2). In this case, each item to be sorted is placed in a carrier on a moving belt or chain and at the loading station a series of cams attached to this carrier is set up to represent the key of the item. The carrier then passes along over a number

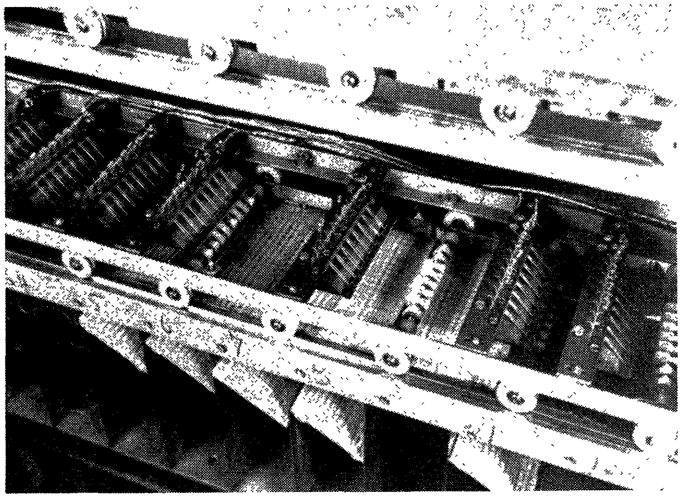
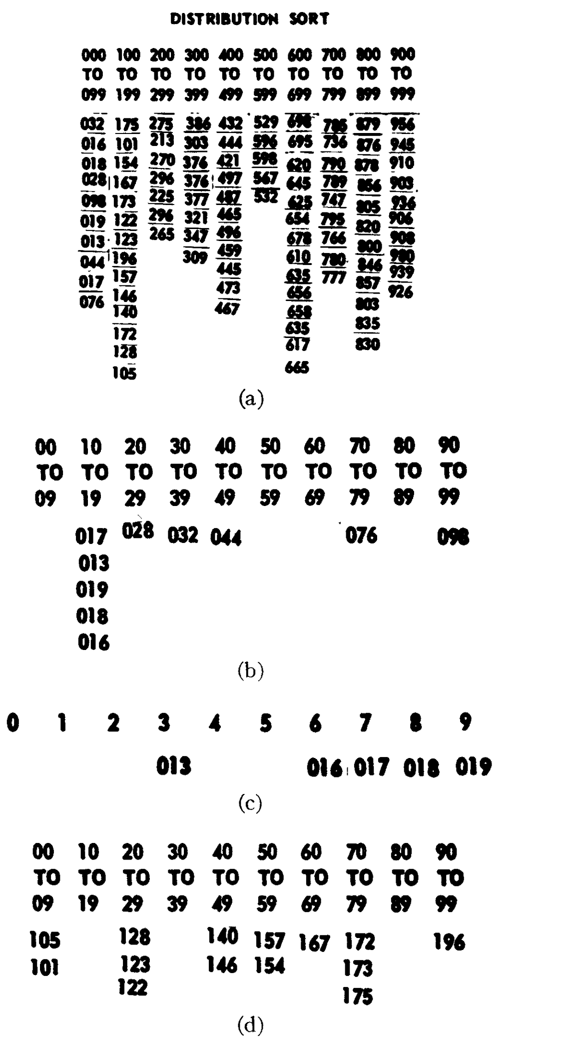


Fig. 2—Rabinow belt sorter at the Bureau of the Census.



(a) Keys distributed into "hundred" blocks (first step).
(b) The first "hundred" block distributed into "ten" blocks (first step of second round).
(c) The second "ten" block of the first "hundred" block distributed into "unit" blocks (second step of third round).
(d) The second "hundred" block distributed into "ten" blocks (second step of second round).

Fig. 3—Distribution sort.

of pockets. At each pocket, a series of two-way switches is actuated or not by the cams. When the cams reach a pocket having a similar setting in its interrogating switches, a solenoid unloads the item from the carrier into the pocket. Systems of this kind can be made with very large numbers of pockets and are suitable for mail applications and for census applications.

The theory of the distribution sort can best be illustrated by a series of diagrams [Fig. 3(a) to (d)].

Digital Systems

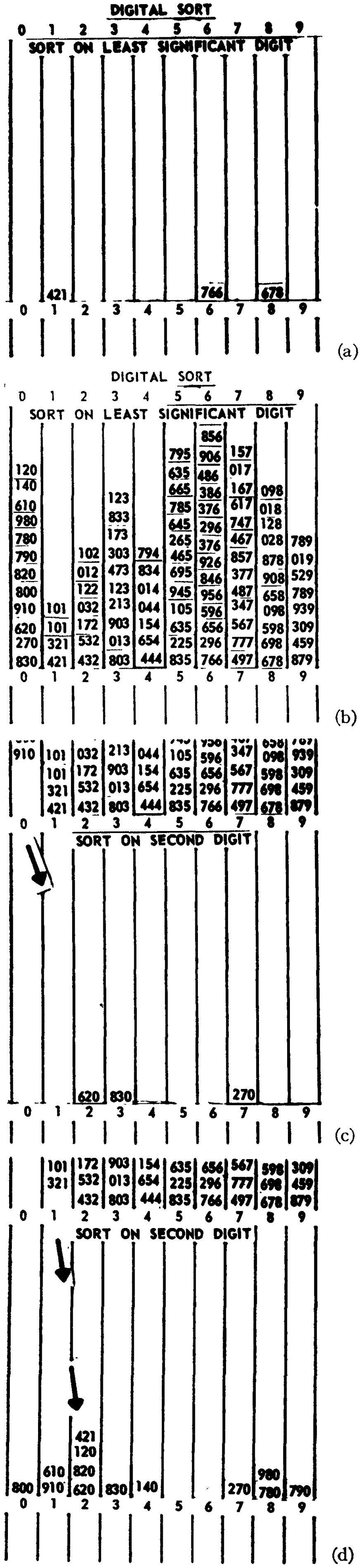
The theory of digital systems can be seen from Fig. 4(a) to (g), p. 42. The new Underwood [7] Rapid-Sort (Fig. 5, p. 43) is a device for digital (or distribution) sorting items like checks, sales slips, or invoices, which are not machine-coded. Each item in turn appears at the top of a pile and can be read by the operator who then presses a typewriter key corresponding to the digit or character against which sorting is being done. The item is then rapidly transported into the corresponding pocket by a series of driven and idler rolls. As soon as the item has passed through the solenoid operated gate, the keyboard is unlocked and the key for the next item can be pressed. Up to 50 pockets are available on this machine.

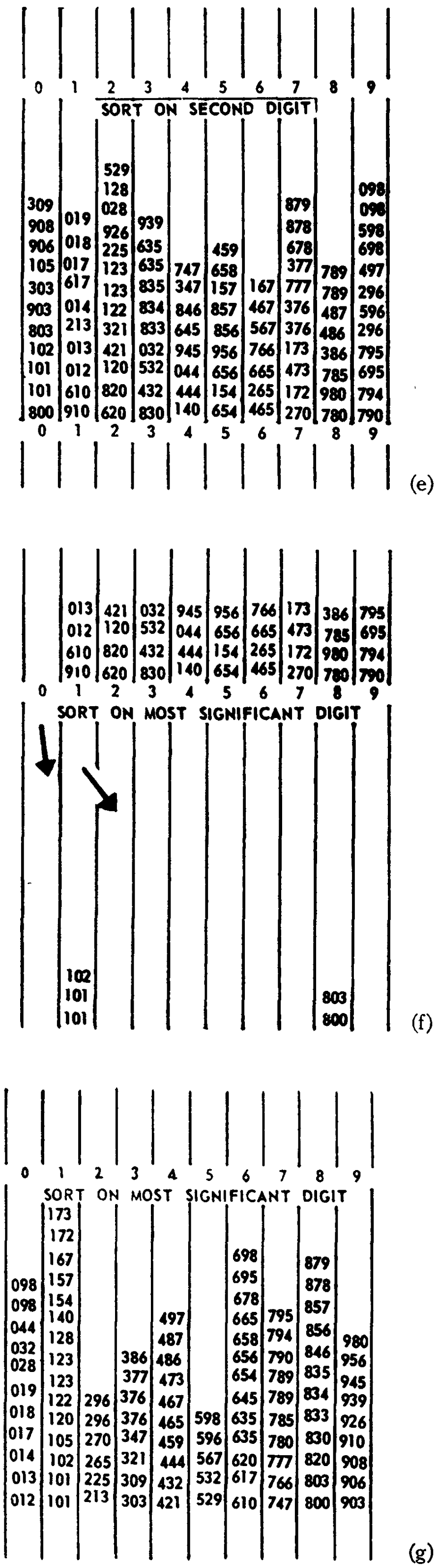
Card sorters in general operate on the digital system. The familiar IBM [8], Remington Rand [9], and Underwood-Samas [10] sorters all operate on generally the same principle. The mechanism of the IBM sorter (Fig. 6) is the easiest to describe and is shown in Fig. 7(a) and (b). At the reading station the solenoid opens a gap between two of the chute blades as the card is fed forward. Thereafter the card is guided by the chute blades into the correct pocket. In the others, cards are guided into pockets by gates. The correct delay between reading and gate opening is derived from a mechanical disc and bail memory.

The Eastman Kodak Minicard system (Fig. 8) has a digital sorter with a circular transport system and is arranged so that cards can be automatically fed into and out of the pockets. No manual operation is therefore required between passes.

Having studied gamesmanship under Stephen Potter [12], I was naturally constrained to turn to Jack Potter [13], for instruction in sortsmanship. He is building a tape sorter which is exactly analogous to the standard card sorter. By making use of a single drive system for all the tape mechanisms, and by using other shortcuts, he hopes to be able to produce a ten-tape sorter for something less than $10,000. The logical system in this device can, of course, be extremely simple since all it has to do is to count to the correct digit position in the key, detect the digit and transfer the item to the tape corresponding to that digit. The information from the 10 tapes in turn is then transferred to the input tape and the operation repeated for the next digit.

Presumably the Marchant Tape Information Processor [14] now being developed with quick change tape magazines will provide comparable results.





(a) Begin sorting on least significant digit (beginning of first round).
(b) Finish sorting on least significant digit (end of first round).
(c) Sort "zero" column by second digit (beginning of first step of second round).
(d) Begin sorting "one" column by second digit (beginning of second step of second round).
(e) Complete sorting on second digit (end of second round).
(f) Sort "zero" column by most significant digit (first step of third round).
(g) Complete sort on most significant digit (end of third round—final step).

Fig. 4—Digital sort.

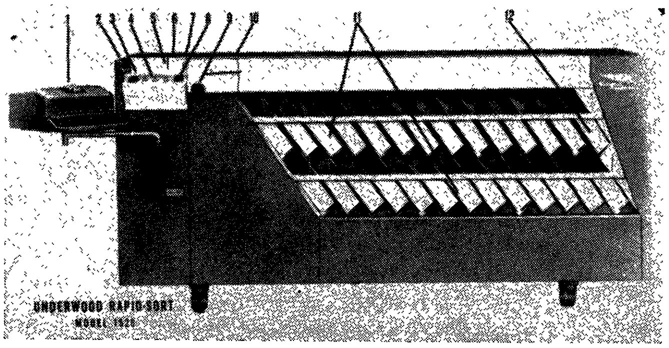


Fig. 5—Underwood Rapid-sort.

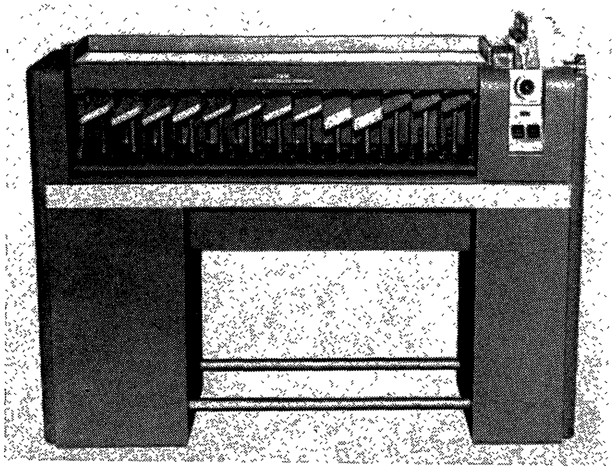
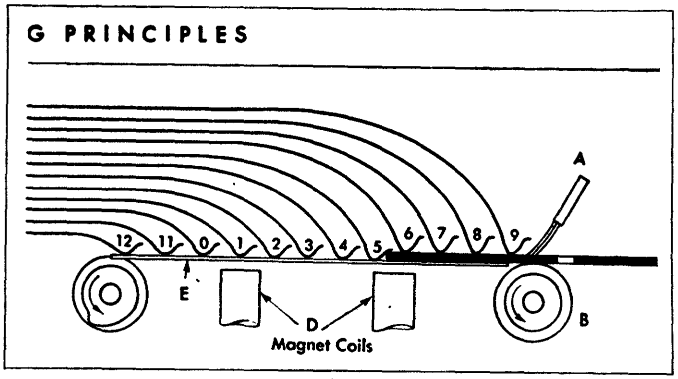
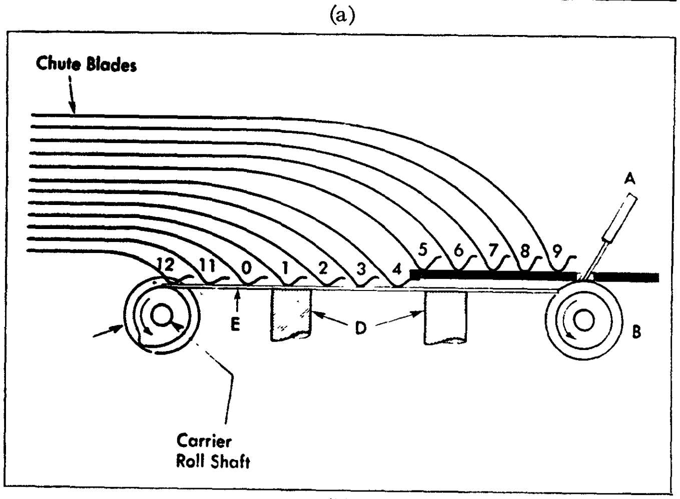


Fig. 6—IBM card sorter model 82.



(a)



(b)

(a) The card is fed forward and pushes up the chute blades.
(b) The brush A senses the code hole and makes the solenoid let the remaining chute blades down so that the card can enter at this point.

Fig. 7—Operating principles of IBM card sorters.

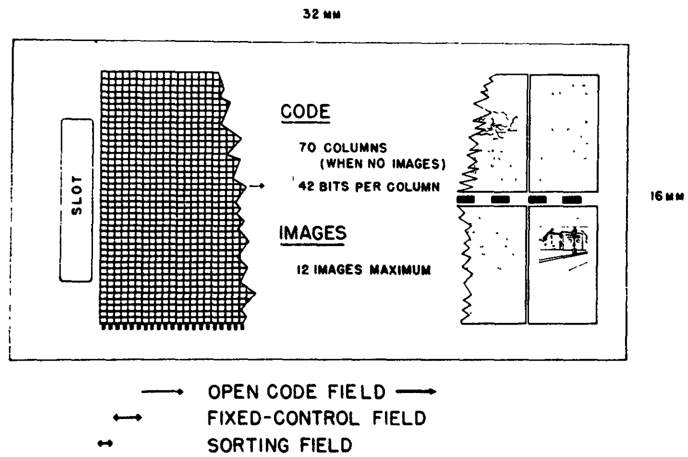


Fig. 8—Eastman Kodak Minicard format.

In their check-handling system for banks, International Telemeter [15] attach a heat-sensitive punched tape to the edge of each check. From then on the checks can be handled in decks on a sort of double skewer and independently on nearly standard punched tape handling equipment. In sorting, the key is read from the holes in the punched tape and when the check reaches the correct position a paddle wheel moves up one notch and removes the check from the conveyor (Fig. 9).

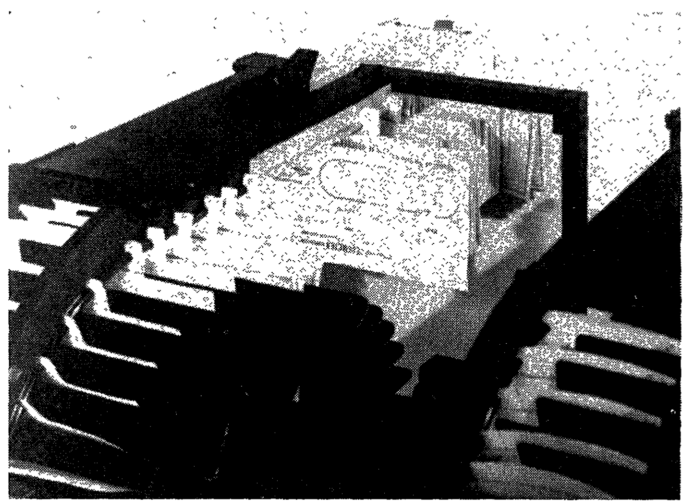
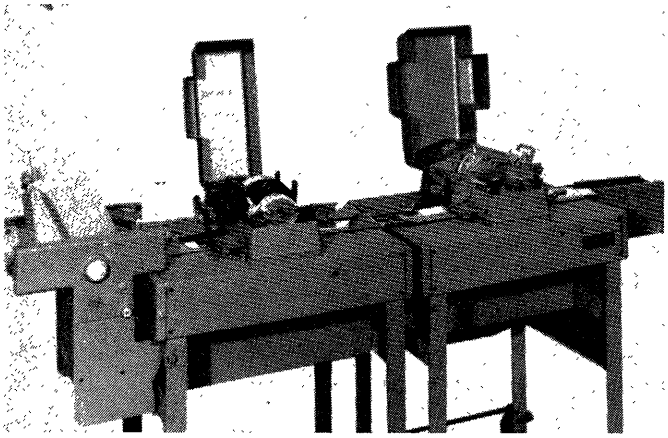


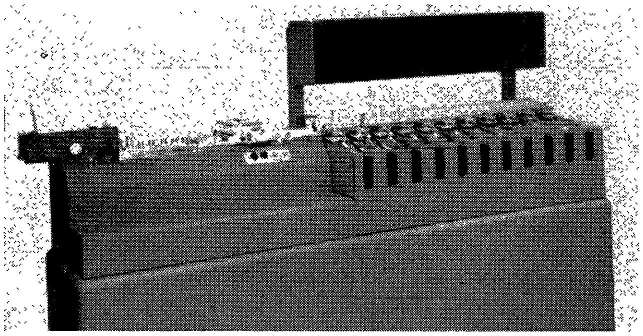
Fig. 9—International Telemeter check sorter.

In the Burroughs [16] Todd [17] Addressograph-Multigraph [18] system for handling checks, the key and other information is contained in dots on the check printed in fluorescent ink [Fig. 10(a)]. The checks are physically sorted by the Todd sorter shown in Fig. 10(b) to (g), p. 44.

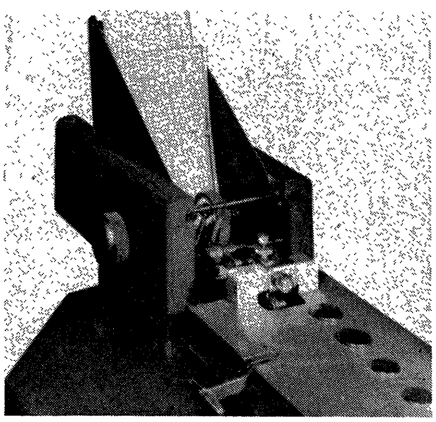
Another check sorter, which is also applicable to other documents, is made by Halm Instrument Co. [19] and used in conjunction with Intelligent Machines Research Corp. equipment, among others. The Halm sorter [Fig. 11(a) and (b), p. 45] is designed in such a way that the check is never gripped by anything more violent than a suction cup. The idea is that it shall do no more harm to the checks than occasionally throwing them on the floor.



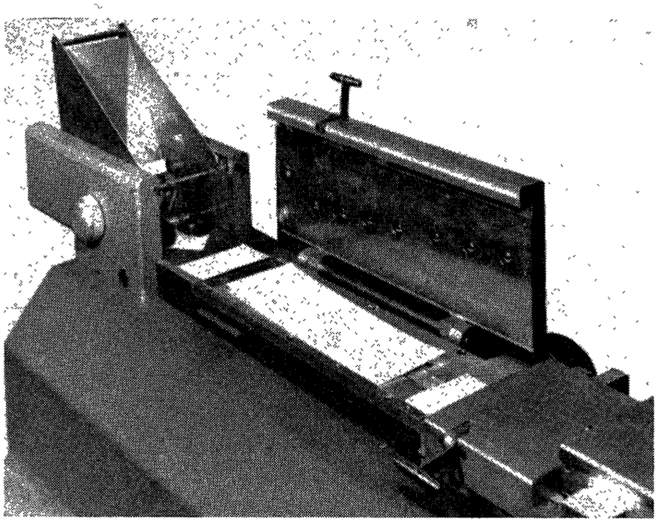
(a)



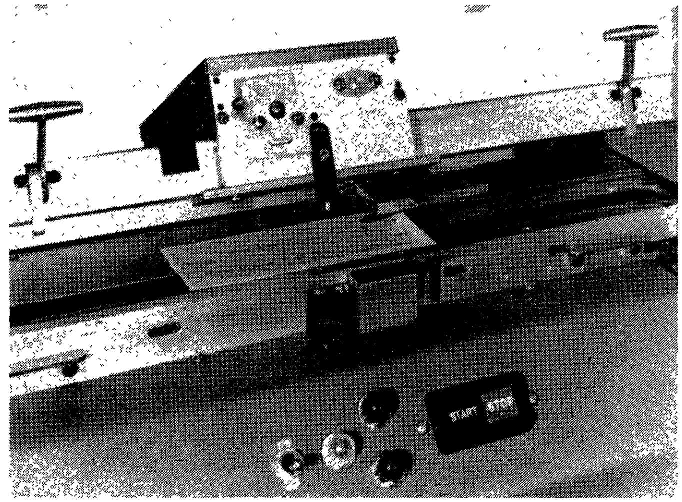
(b)



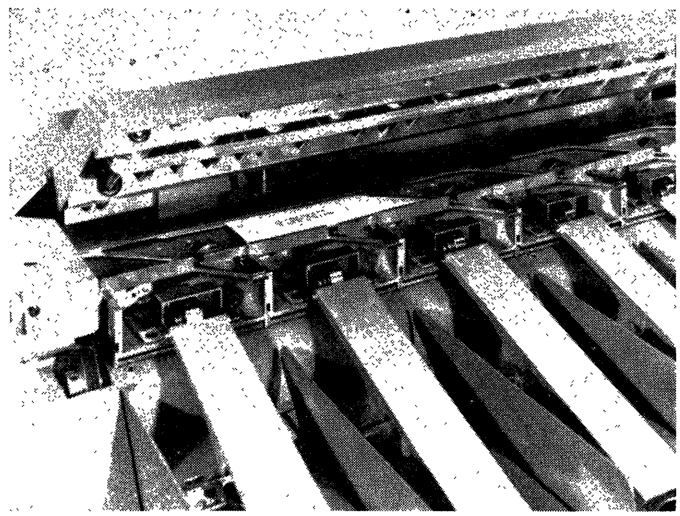
(c)



(d)



(e)

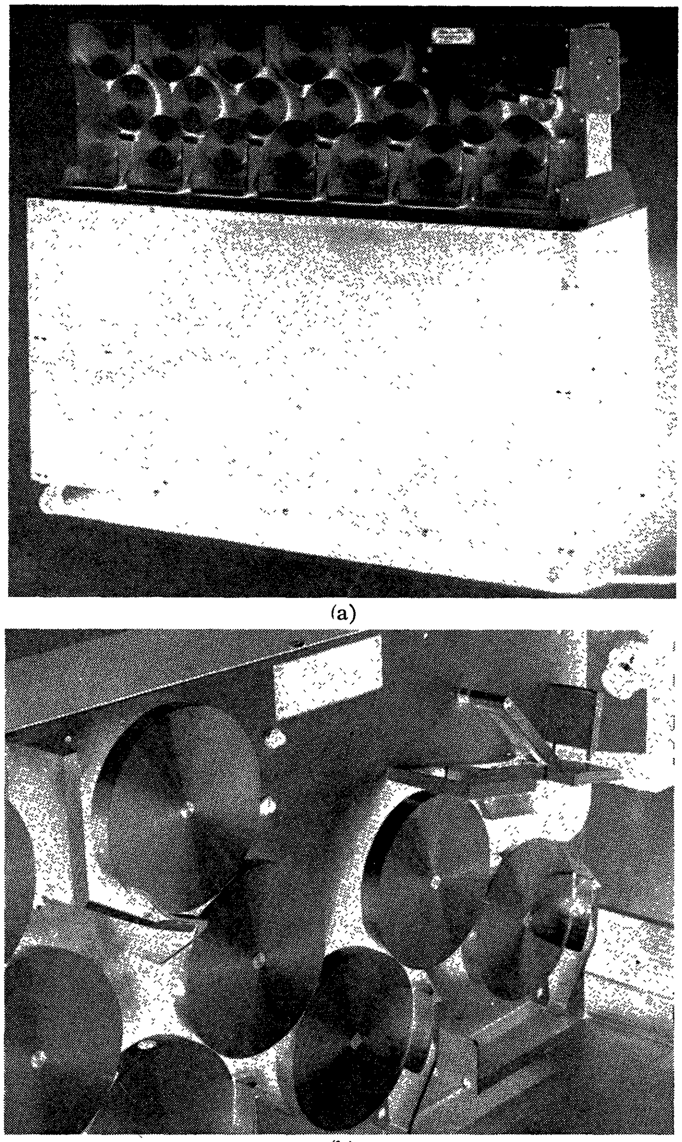


(f)

(g)

(a) Check imprinter.
(b) Over-all view of sorter.
(c) Checks being fed from the top of the pile.
(d) Skewed belt and balls align check with edge of track.
(e) Reading section open showing check entering chute blades.
(f) Distribution section open showing check entering turning bar.
(g) Check entering pocket upside down to maintain original order.

Fig. 10—Todd check-sorting equipment using fluorescent ink dot code.



(a)

(b)

(a) Twelve pocket pneumatic sorter.
(b) Detail showing reading head and transfer discs with suction ports.

Fig. 11—Halm sorter.

ERMA was somewhat coy about her sorting equipment and all I have is a photograph (Fig. 12) taken in the dark. I was, however, told that her sorting mechanism involved electrical, mechanical, and air-operated devices. No doubt in the immortal words of T. S. Eliot she "gives promise of pneumatic bliss."

The Automatic Message Accounting System (Fig. 13) devised by the Bell Laboratories [20, 21] for the telephone companies is one of the few examples of punched tape sorting equipment. This operates on a decimal-digital system.

The Sperry Rand File Computer [43] is primarily an inventory machine but has been built to have as much flexibility as possible. Its coding distributor and large capacity drum memory allow the machine to be programmed to perform a digital sort on two decimal digits at a time thus halving the number of passes required with a normal digital sort.

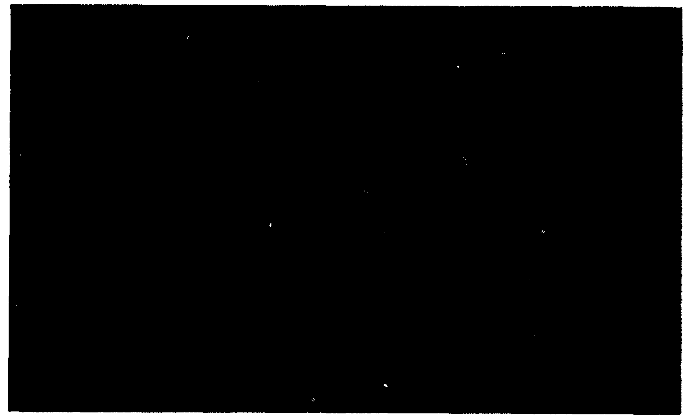


Fig. 12—ERMA photographed in the dark.

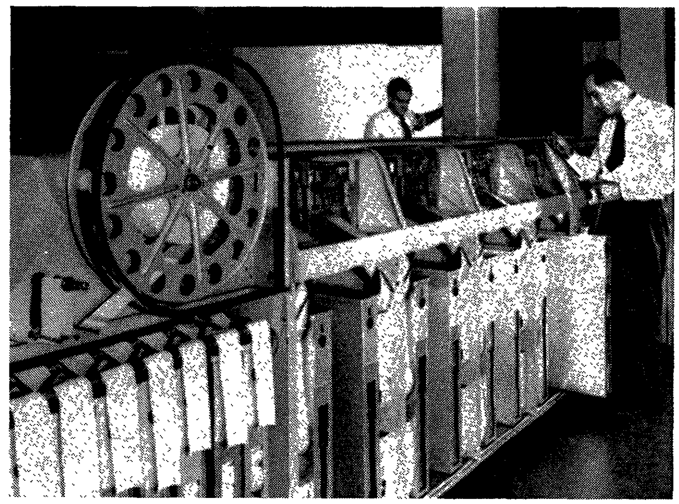


Fig. 13—Bell Automatic Message Accounting System.

Mesh or Merge Systems

These systems are the commonest with magnetic tape equipment. The theory of two of them is explained in Fig. 14(a) to (m), pp. 46 and 47, and Fig. 15(a) to (g), pp. 48 and 49.

The Underwood Elecom File Processor [22–24] uses a two-way merge sorting system.

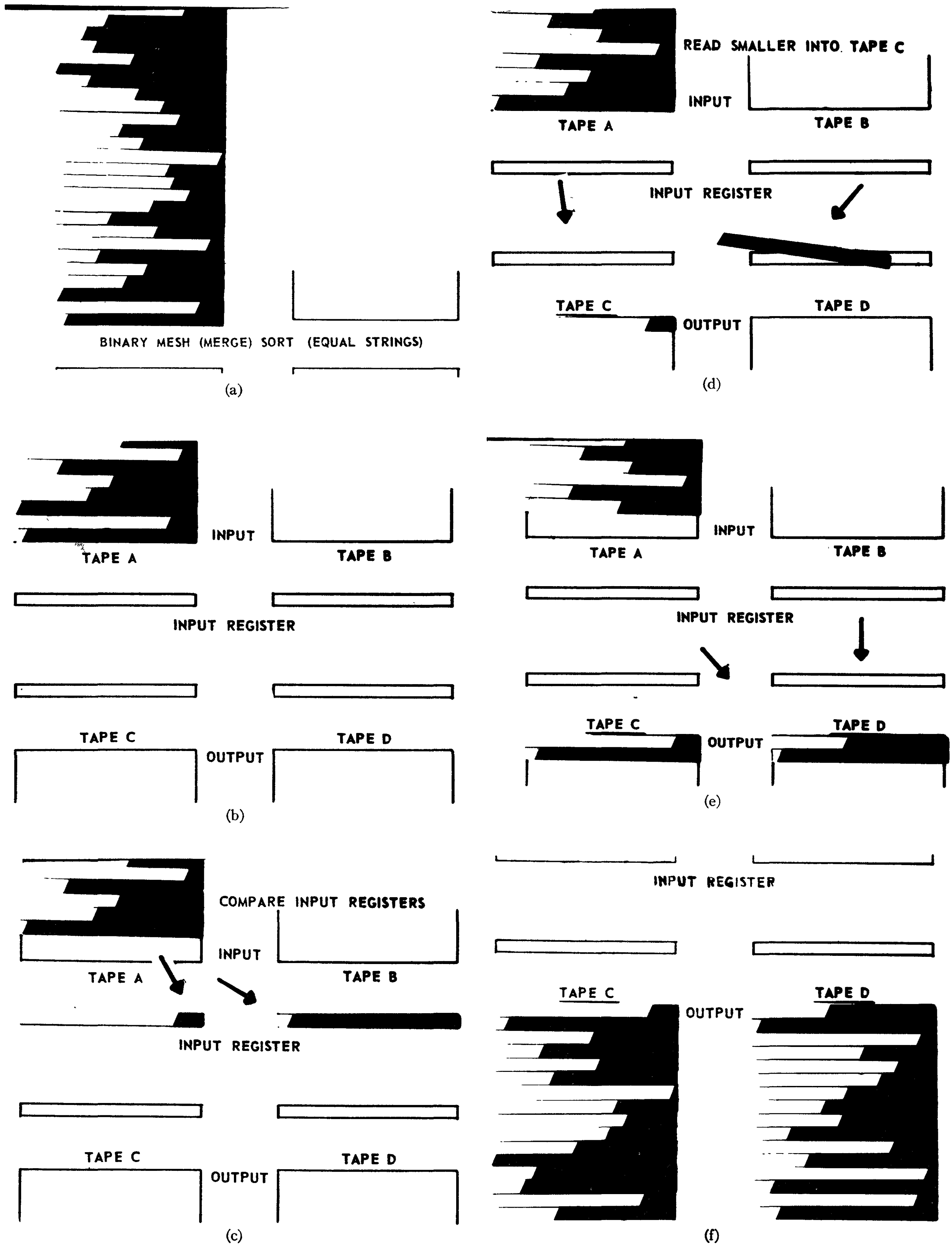
The RCA Bizmac Sorter [25, 26] uses a two-way merge system.

Univac [27–31] at first used two-way merge sorting and the Univac sort generator was the first major automatic routine to be completed. It has been in use since 1951. A three-way sort routine has now been devised for Univac. Also available is a more sophisticated merge system making good use of the high speed memory and known as the *Arranger* [32].

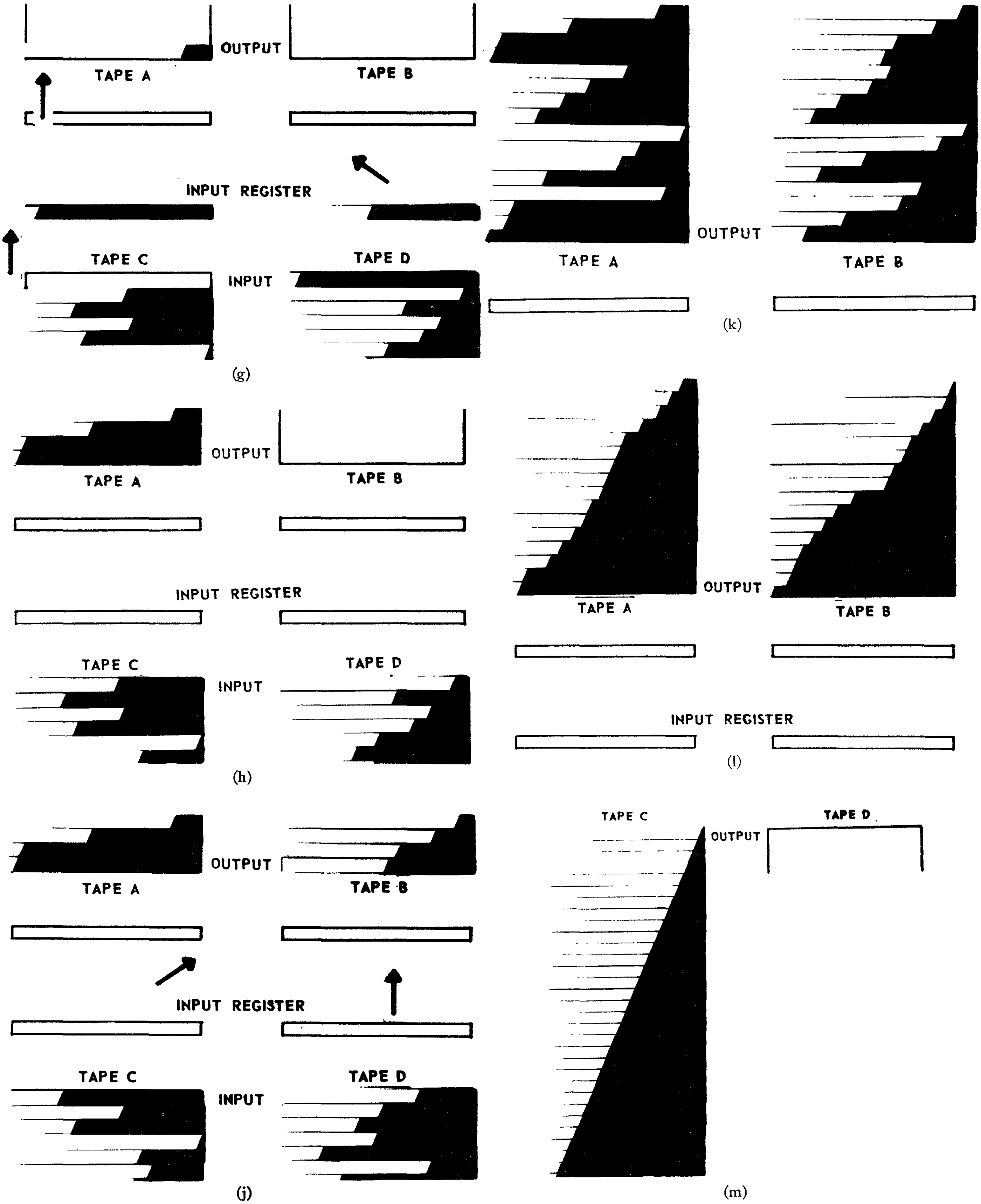
For the IBM 705 [33–36] a sophisticated three-way sort routine is available.

The Datamatic 1000 [37, 38] allows a very fast read in from its wide tape and its sorting capabilities are rumored to exceed those of the other machines.

Time is saved at the beginning of most of these merge-sort routines by reading blocks of items into the internal memory and there sorting them into strings at



(a) Input tape; items represented by strips; the length of the black part representing the key.
(b) Beginning of first pass; input from Tape A; output to tapes B & C; two input registers each holding one item.
(c) Read first two items into input registers.
(d) Compare, read smaller and then larger into Tape C.
(e) Read next two items in ascending order into Tape D.
(f) Repeat till Tapes C & D each contain a series of strings **of two.**



(g) Exchange input and output; read from each input tape into corresponding input register; compare; read smaller into output Tape A; immediately replace with next item from the tape the smaller came from; compare and read smaller into Tape A.

(h) Continue thus until the first two strings of two in Tapes C & D have been merged into a string of 4 in Tape A.

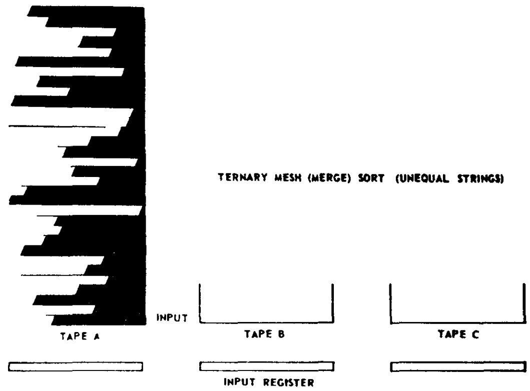
(j) Merge the next two strings of two into a string of 4 in Tape B.

(k) Repeat till all strings of two have been merged into strings of 4 in Tapes A & B.

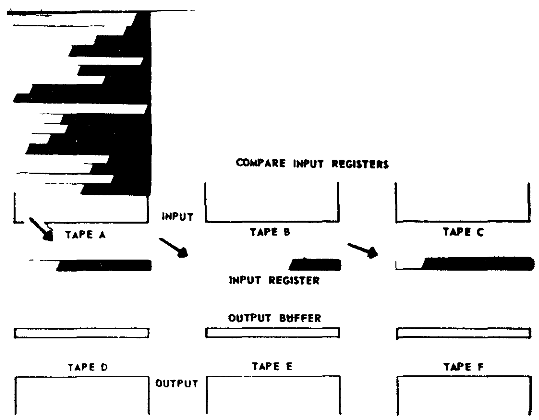
(l) Merge these into strings of 8 in tapes C & D and then into strings of 16 in Tapes A & B.

(m) Merge these into a string of 32 in Tape C.

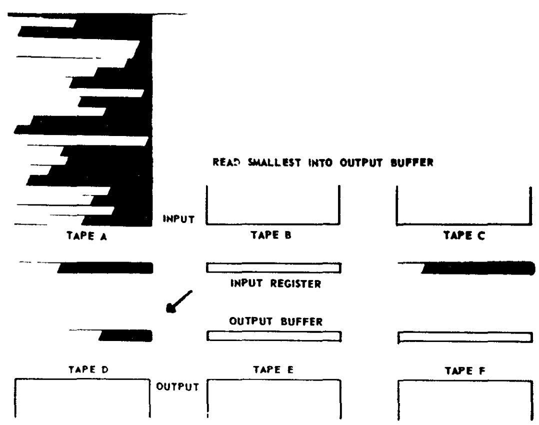
Fig. 14—Binary mesh (merge) sort; equal strings; input in random order.



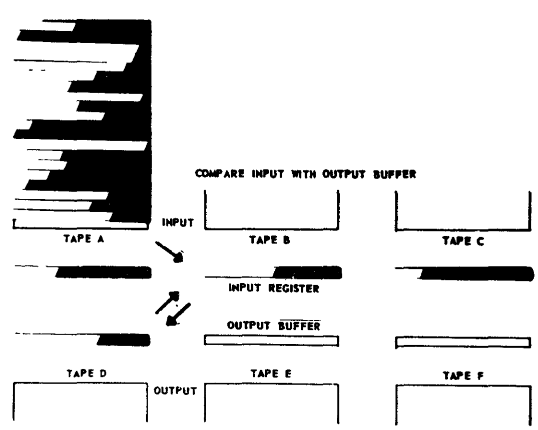
(a)



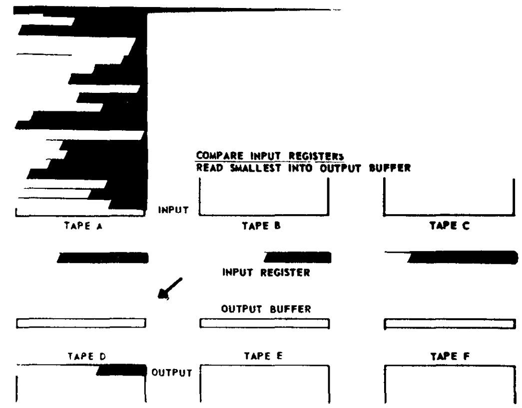
(b)



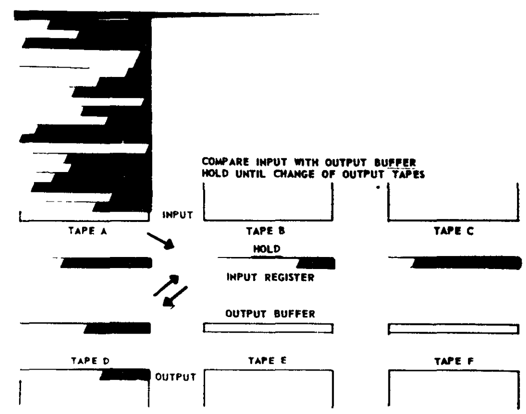
(c)



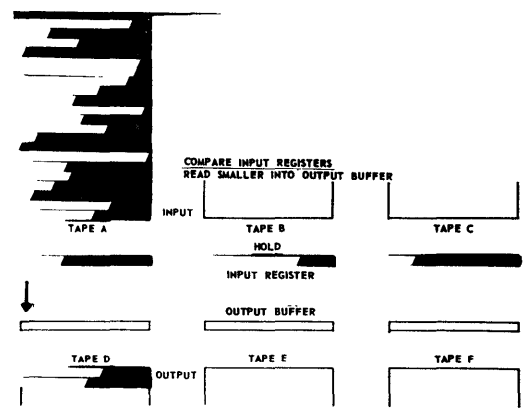
(d)



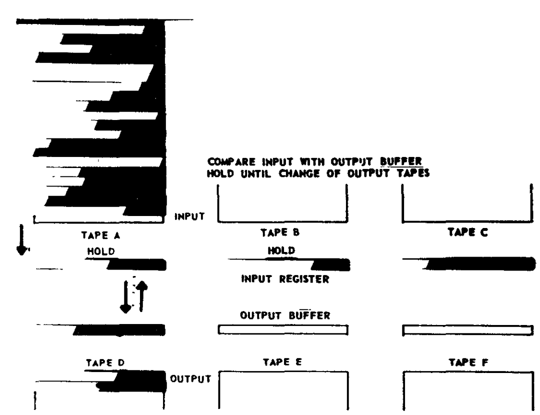
(e)



(f)

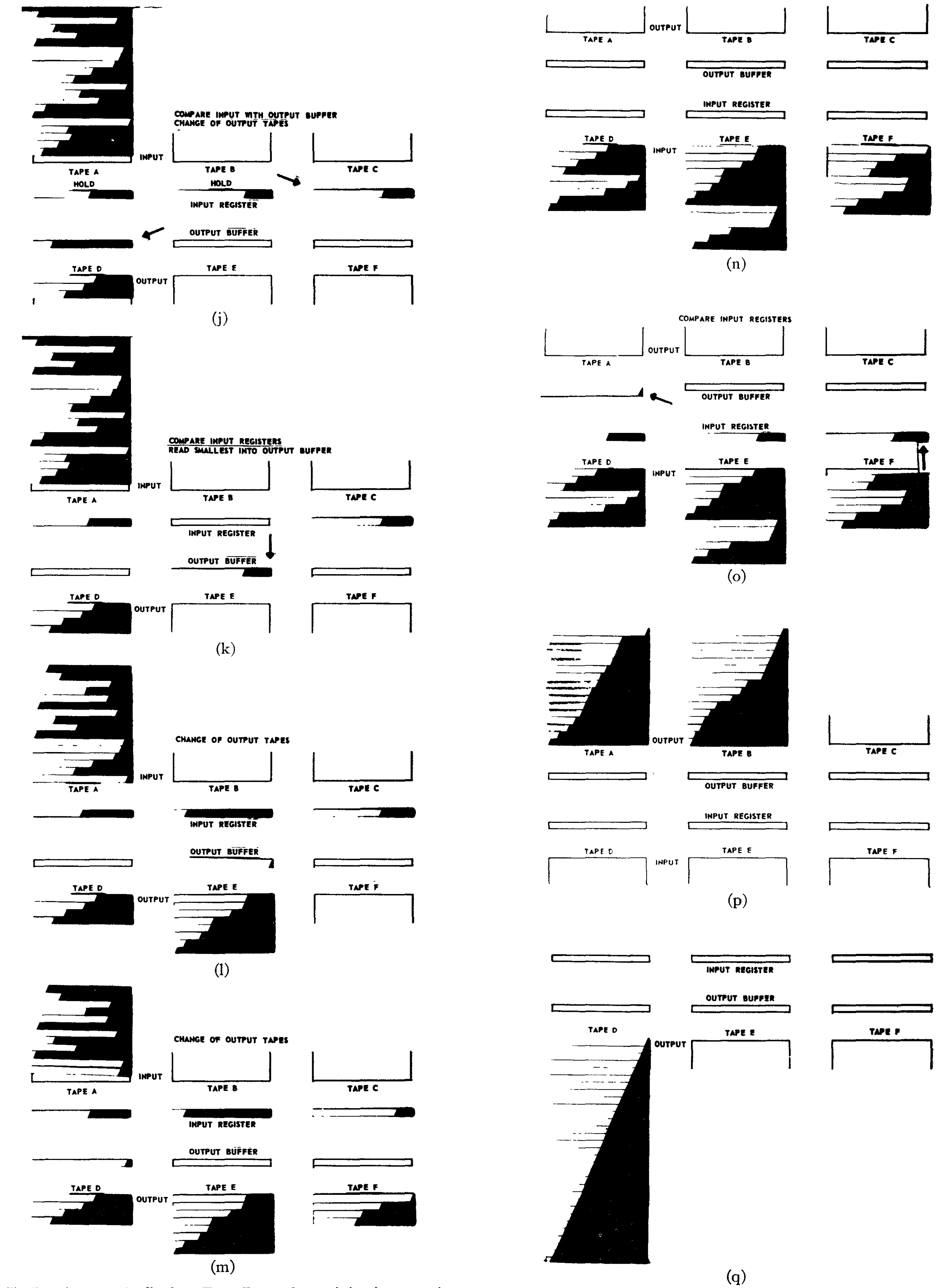


(g)



(h)

(a) Input Tape A; items in random order.
(b) Read first three items into input registers.
(c) Compare A with B and smaller with C; read smaller into output buffer D.
(d) Read next item into empty input register; compare with output buffer D; if larger continue.
(e) Read output buffer into Tape D; compare input registers; read smallest into output buffer D.
(f) Read next item into empty input register; compare with output buffer D; if smaller hold until change of output tapes.
(g) Read output buffer into Tape D; compare remaining input registers; read smaller into output buffer D.
(h) Read next item into empty input register; compare with output buffer D; if smaller hold until change of output tapes.



(j) Read output buffer into Tape D; read remaining input register into output buffer D; read next item into empty input register; compare with output buffer D; if smaller change output tapes.

(k) Compare input registers; read smallest into output buffer E.

(l) Continue in the same way to read a string into Tape E.

(m) Continue to read a string into Tape F.

(n) Continue to read further strings into the output tapes in turn.

(o) Change input to output; read first item of each input tape into corresponding input register; compare and read smallest into output buffer A.

(p) Continue to merge Tapes D, E, & F into strings in Tapes A, B, & C.

(q) Change input to output and merge these strings into one string in Tape D.

Fig. 15—Ternary mesh (merge) sort; unequal strings; random input.

high speed. This takes the place of the first few passes. Another method of time saving, used for example in the 705 routine mentioned, is to read the information from the input tape into the internal memory in large blocks. This saves much of the time used normally in starting and stopping the tape. The *Arranger* [32] routine already mentioned makes use of the internal memory to convert the input into variable length strings, any of which may contain more information than the capacity of the internal memory. In merging it effectively uses parts of the high speed memory as large input registers and output buffers to reduce the number of passes needed.

The National Cash Register Co. [39] designed a processor, the Model 303, which incorporates a built-in sort command to sort the contents of a full drum channel and a built-in merge command to merge two sorted channels.

Address Calculation

Address Calculation, a relative newcomer among sorting routines, was lately described by Mr. Isaac of SRI [40]. In a way the system is similar to that used in the Savasort [40] hand sorting system shown in Fig. 16.

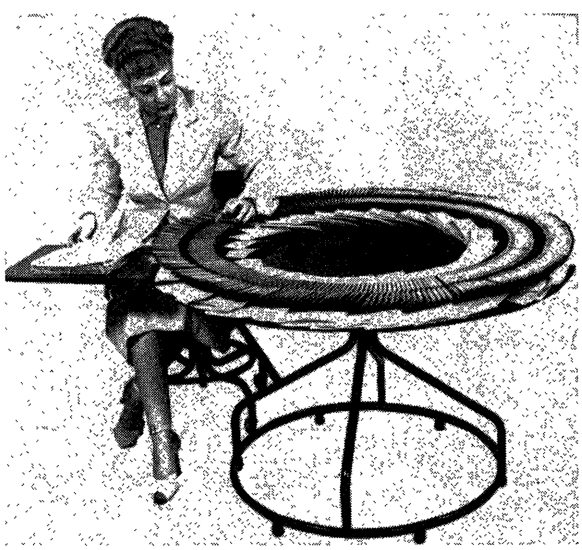


Fig. 16—Savasort manual sorting device.

In this manual system the operator makes a guess at the approximate position in the file, turns the file to that point, and then checks the exact position in which to insert the item. The address calculation method for rapid access memories requires some knowledge of the characteristics of the keys of the items to be sorted. From the information, an approximate relation is established between the key and the address in which it has to be inserted. Suppose that 50 items are to be sorted into 100 pockets and that their keys are random numbers between 0 and 999. Suppose the pockets are numbered 0 to 99. A good approximation at sorting would be obtained by dividing the key by 10 and inserting in the pocket corresponding to the answer. Fig. 17(a) to (h) shows the theory and the routine for making slight corrections, if the pocket happens to be full. This system is only efficient if the number of pockets available is nearly twice as great as the number of items. Otherwise, too much comparing and moving is needed.

Another insertion is used by Monroe [42] in their Monrobot under some conditions. They program the search so that the required address on the drum is traced as being first in one half, then in one half of that half, and so on.

Selection Systems

The theory of this technique is shown in Fig. 18(a) to (j), p. 52.

A variety of this technique has been described by Mr. Friend [44] under the name "Nth Degree Selection."

A similar method has been used by the ERA group of Sperry Rand [43].

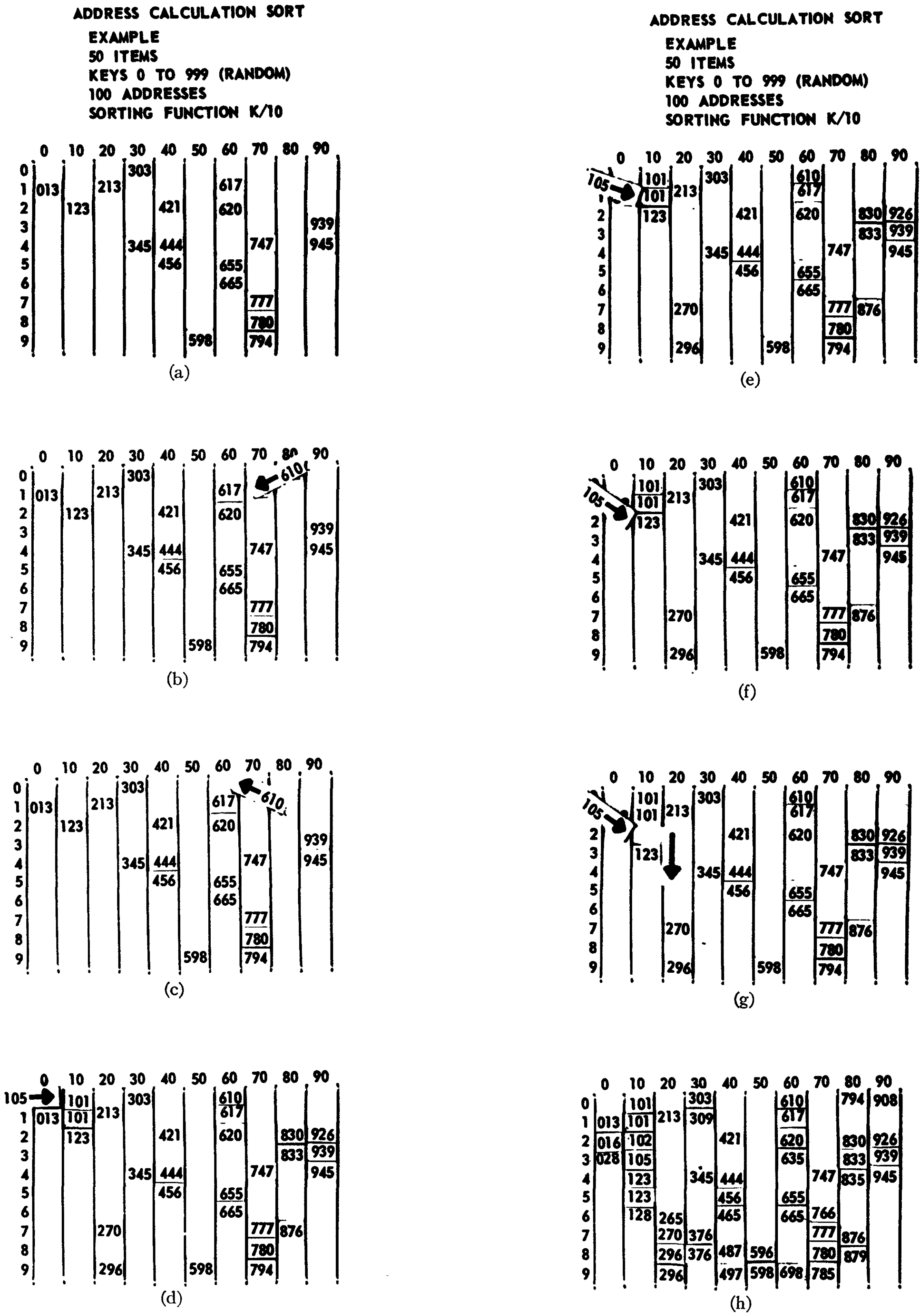
The system is used by The Laboratory for Electronics in Diana (of the Chase National Bank) [45, 46]. Their system [Fig. 19(a) and (b), p. 53] uses slow-speed, high capacity drums and sorting is not normally used until a printed output is required. By means of a selection system, after an initial run of a few minutes, the Diana sorter produces sorted information at the rate of $2\frac{1}{2}$ items per second, which is enough to match the speed of a line-at-a-time printer. Sorting on the Diana computer takes twice as long.

Exchange Systems

This method has very simple logic but requires too large a number of operations to be of much use generally. Items in addresses 1 and 2 are compared and the one with a smaller key read back into address 1. The key of the remaining one is compared with that of the item in address 3 and the smaller read into address 2 and so on. Walter Soden [47] has suggested a modification which he claims reduces the number of operations.

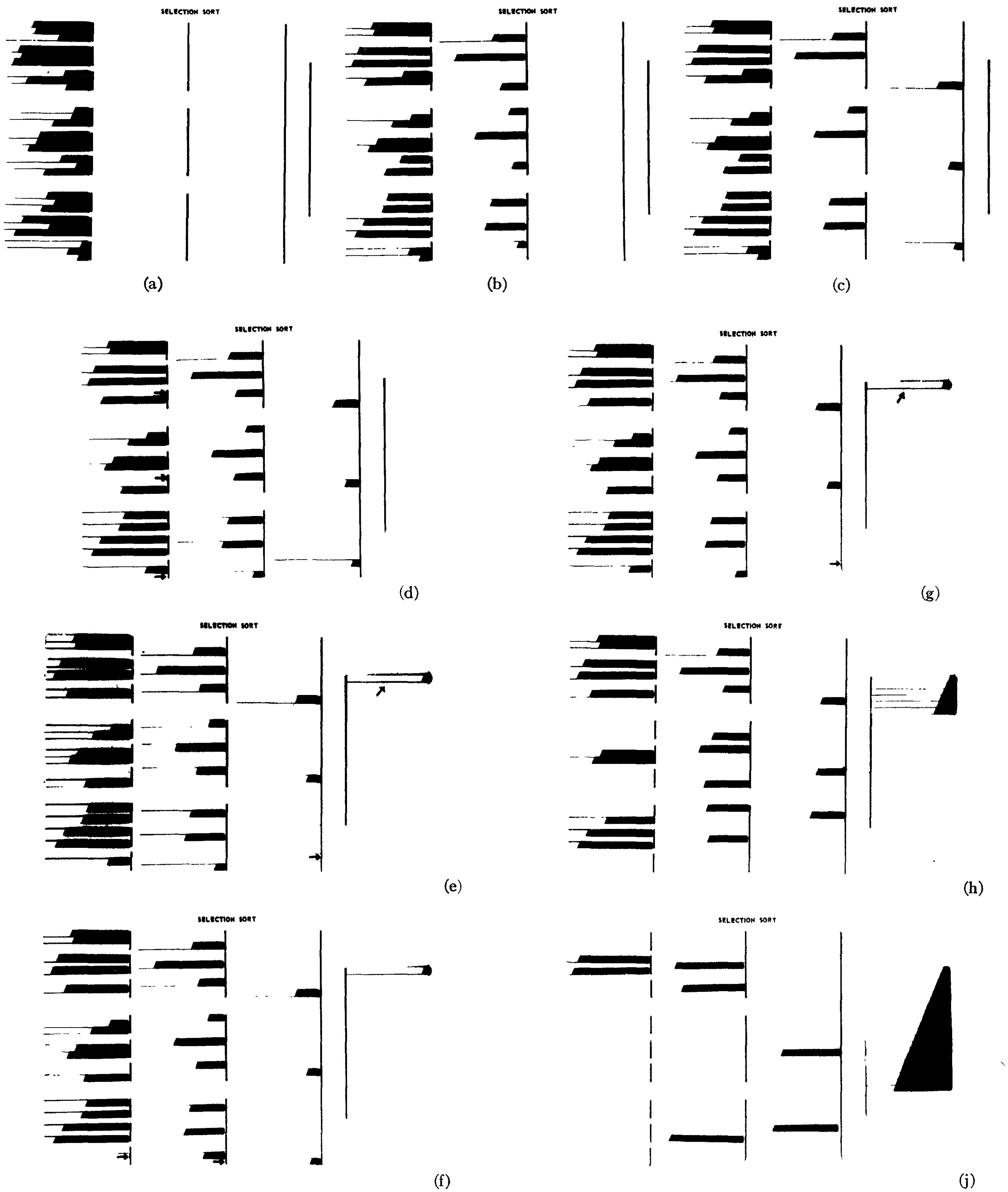
Avoid Sorting

The Clevite [48, 49] "TapeDRUM" [Fig. 20(a) and (b), p. 53] which effectively gives magnetic drum type access to a page at a time of information carried on a wide tape, can be used by various means to avoid sorting. For example, where information about transactions is coming into a system and has to be used to update a large file, the file can be held on the tape of a "TapeDRUM" and the incoming items can be temporarily stored on a normal magnetic drum with enough capacity to hold the number of items which comes in in the time required for the "TapeDRUM" to go through the whole length of the tape. In use, the "TapeDRUM" scans a page at a time and wherever the key of an item on the high-speed drum corresponds with the key of an



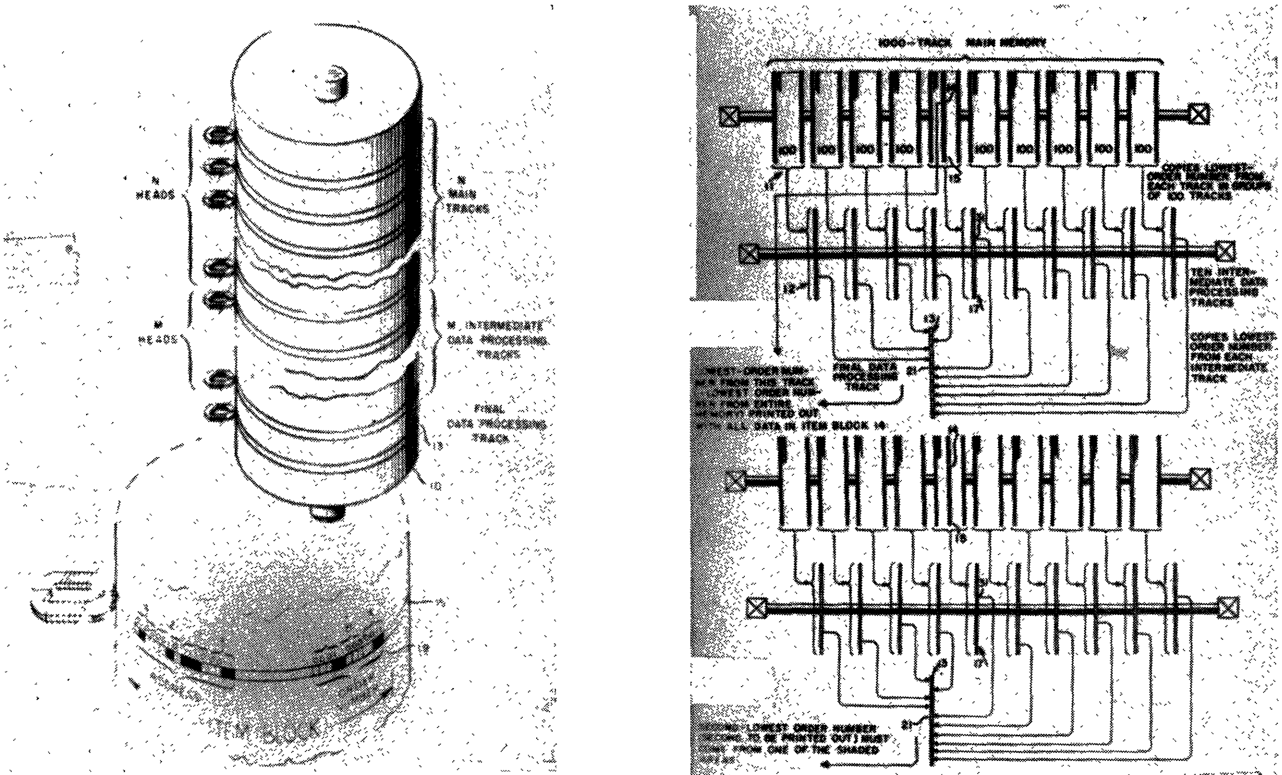
(a) Read into calculated addresses.
(b) Address full; compare with item in address; if smaller move down one address.
(c) If empty read in.
(d) Address full; compare; if larger move up one address.
(e) Address full; compare; if larger move up one address.
(f) Address full; compare; if smaller move item in address up one.
(g) Read into empty address.
(h) Fifty items inserted in 100 addresses.

Fig. 17—Address calculation sort.



(a) By a series of comparisons, select smallest key in each track (one revolution).
(b) Read these items with the number of the tape they came from into the three partial sort tracks.
(c) Select item with smallest key in each partial sort track and read item into final sort track.
(d) Replace each item removed from the partial sort tracks by the next smallest in the track it came from.
(e) Select the smallest in the final sort track and read it into the output.
(f) Replace this by the smallest in the partial sort track it came from; replace the latter by the smallest left in the track it came from.
(g) Read out the smallest in the final sort track; replace as before.
(h) Continue selecting and replacing.
(j) Continue as before.

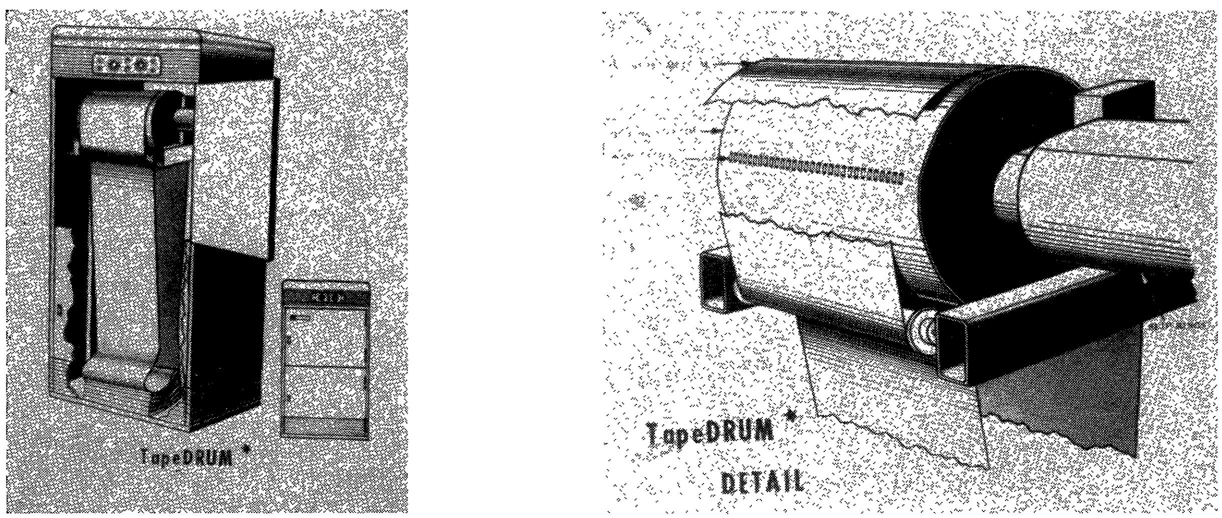
Fig. 18—Selection sort—example: 9 tracks each containing 3 items are sorted by using 3 partial sort tracks, 1 final sort track and an output.



(a) (b)

(a) Arrangement of tracks on drums.
(b) Schematic of selection sorting routine.

Fig. 19—Diana system of selection sorting.



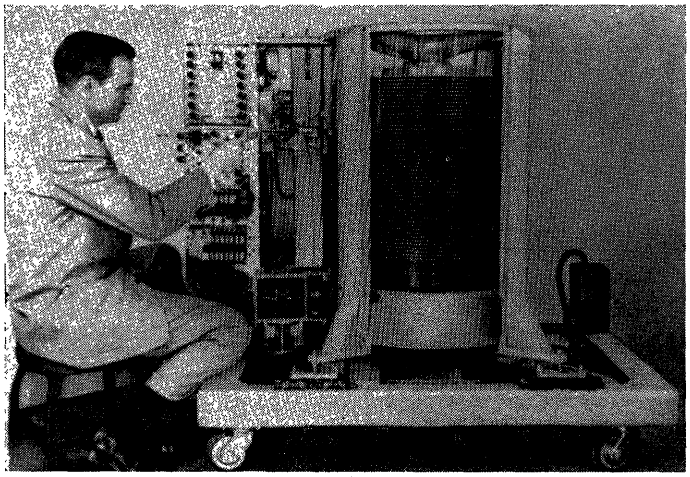
(a) (b)

(a) General arrangement.
(b) Detail of drum and wide tape.

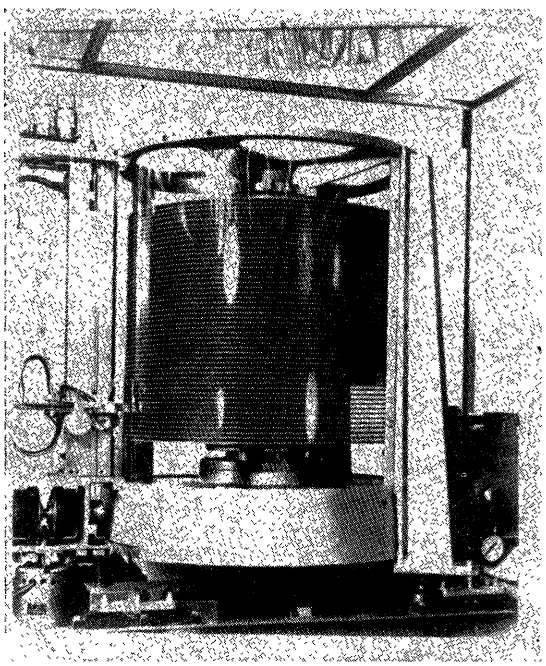
Fig. 20—Clevite TapeDRUM.

item on the "TapeDRUM," that item is transferred to the "TapeDRUM." If the final information has to be sorted by more than one key, two "TapeDRUMS" can be used.

Examples of large capacity rapid access memories other than that of Diana are the IBM multiple magnetic disk memory [50, 51] usually known as the Juke Box [Fig. 21(a) and (b), p. 54] and the Potter [13] and Telecomputing [52] rapid access magnetic tape memories. All of these are especially suitable for systems in which the information has to be arranged always according to one key. The Photoscopic [15, 53] information storage system (Fig. 22, p. 54) being developed by International Telemeter operates on different princi-



(a)



(b)

(a) The read-write head is raised or lowered and inserted mechanically.
(b) The disks which hold up to 5 million characters.

Fig. 21—IBM 305 disk (Juke Box) memory.

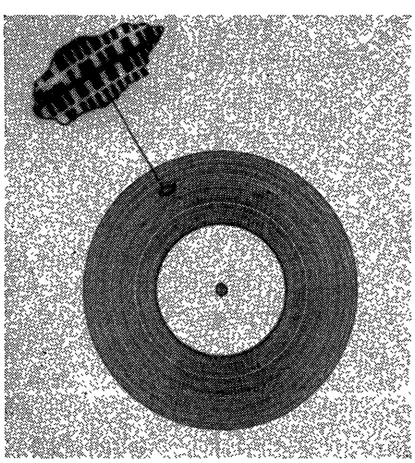


Fig. 22—International Telemeter Photoscopic Memory.

ples. In this case, the information is contained in a photographic film attached to a rotating glass disk. The access time in this system is much smaller than in the other systems and an access time of 50 milliseconds to any word in a 20 million bit store is being talked of. The main difference between this system and the others is that, being photographic, the information is not erasable. There are, however, a number of applications where this is not very important. Examples are a mail-order catalog, a dictionary, and the information required for sorting mail down to mailmen's routes. If a relatively small number of corrections and additions are required, this can be achieved by using a magnetic drum memory in parallel, having a capacity some small proportion of that of the Photoscopic store.

Comparison of Systems

A great deal has been written about methods of comparing the speeds of various sorting routines [54–59] but in a paper of this kind all I can do is to provide a bibliography and extract some generalizations. In choosing between various sorting systems, it appears to be unsafe to reach a decision without first examining all the requirements, all the available equipment and several possible systems. It can be said in general, however, that distribution systems are likely to be faster when the number of possible keys is small compared with the number of items; that digital systems are good when the number of possible keys is about equal to the number of items and where very simple logic is a help; and that merge systems are suitable when the number of possible keys is large compared with the number of items, where more complicated logic can be used to speed up the process, or where items are partially ordered to start with.

Actual sorting time cannot be computed until information is available on the number of words per item, the number of digits in the key, the number of items, the speeds of the various parts of the sorting mechanism, and the details of other clerical functions which can be combined with sorting.

To achieve economy in a clerical system, it is necessary that the various parts be in balance; that is to say that no one function such as tape handling acts as a bottleneck and prevents other parts of the equipment from operating at full capacity. In line with this reasoning, many of the recent improvements in sorting have been brought about by increasing the speed at which information can be read in and out of tapes (vide the Datamatic 1000 [37], which uses 3 inch wide tapes traveling at 100 inches a second read at a rate of 56,000 characters per second). The use of ample buffer storage which permits reading, computing, and writing at the same time is now common practice.

In some cases, economy is achieved by separating sorting and computing functions. The commonest case is where the information is on cards which can be sorted very economically. In many clerical systems involving magnetic tape files which have to be kept up to date a very large proportion of the tape handling time is taken up in sorting input transactions and merging these with the main file. In a typical case 5,000 items out of a file of 100,000 may have to be brought up to date. If this were done in a computer, most of the computer operating time would be spent searching tape and only a small part of the time in computing. One approach to this problem, used in Bizmac [26], Miniac [14], and Elecom [24], is to have a separate relatively unsophisticated fixed program selecting, merging, and sorting machine which prepares short tapes for the main computer's concentrated attention. The logic of selecting, merging, and two-way merge sorting is very similar. Of course, this choice would not be economic if it involved idle computer time. This again emphasizes the necessity for analyzing the job and using balanced machine elements. Diana [46] has a separate sorter with some additional parallel computing components which allow a sorting output twice as fast as that from the computer.

I have touched lightly on many facets of a wide field. I shall have accomplished my purpose if I have sorted and put in logical order the gist of the great volume of material which has been generated about sorting over the past few years.

Bibliography

[1] Electronic Recording Machine Accounting, Stanford Research Institute, Menlo Park, Calif. *Fortune* (October, 1955) p. 130.
[2] Project under Dr. M. Levy, Canadian General Post Office, Ottawa, Canada.
[3] Intertype Corporation, 360 Furman St., Brooklyn 2, N. Y.
[4] Mergenthaler Linotype Company, Inc., 29 Ryerson St., Brooklyn 5, N. Y.
[5] Rabinow, Dr. Jacob, Rabinow Engineering Company, 7212 New Hampshire Ave., Washington, D. C.
[6] Conveyor belt sorter, Technical Report No. TR-3, Diamond Ordnance Fuze Laboratories, Ordnance Corps., Dept. of the Army, Washington, D. C., October 20, 1953.
[7] Underwood Corporation Accounting Machine Division, 1 Park Ave., New York 16, N. Y.
[8] IBM Electric punched card accounting machines. Principles of operation. Sorters type 82, 80, and 75, International Business Machines Corporation, 590 Madison Ave., New York, 22 N.Y.
[9] Type 320 sorter and type 420 (electronic) sorter, Remington Rand Division, Sperry Rand Corporation, 315 Fourth Ave., New York 10, N.Y.
[10] Samas Punched Card Division, Underwood Corporation, 1 Park Ave., New York 16, N. Y.
[11] Tyler, A. W., Myers, W. L., and Kuipers, J. W.,"The application of the Kodak Minicard system to problems of documentation," Eastman Kodak Company, Rochester 4, N. Y. Presented at annual meeting American Documentation Institute, Cleveland, Ohio (November, 1954).
[12] *The Theory and Practice of Gamesmanship*. Potter, Stephen, New York, Henry Holt and Company, Inc.
[13] Potter Instrument Company, 115 Cutter Mill Road, Great Neck, L. I., N. Y.
[14] Marchant Calculators, Inc., Oakland 8, Calif.
[15] International Telemeter Corporation, 2000 Stoner Ave., Los Angeles 25, Calif.
[16] Burroughs Corporation, Research Center, Paoli, Pa.
[17] The Todd Company, Inc., Rochester 3, N. Y.
[18] Addressograph-Multigraph Corporation, 1200 Babbitt Rd., Cleveland 17, Ohio.
[19] Halm Instrument Company, Glen Head, L. I., N. Y.
[20] AMA (reprints from Bell Laboratories Record) Bell Telephone Laboratories, Murray Hill, Summit, N. J.
[21] Meszar, John, "Fundamentals of the AMA System." *Transactions of the American Institute of Electrical Engineers*, Vol. 69 (1950), p. 255.
[22] Electronic Computer Division, Underwood Corporation, 35-10 36th Ave., Long Island City 6, N. Y.
[23] Elecom 125 Electronic Business System, Underwood Corporation (1955).
[24] Elecom File Processor Operating Manual, Underwood Corporation (1955).
[25] RCA Victor Division, Radio Corporation of America, Camden, N. J.
[26] RCA Bizmac Programming Manual (1955).
[27] Univac Division of Sperry Rand Corporation, 315 Fourth Ave., New York 10, N. Y.
[28] Sorting Methods for Univac System, Remington Rand (1954).
[29] Catalog of Univac System Routines, Remington Rand Division of Sperry Rand Corporation (1955).
[30] Master Generating Routine for 2-Way Sorting, Eckert-Mauchly Division, Remington Rand, Inc., 965-1 (September, 1952).
[31] Holberton, F. E., Sorting Rerun Procedure, incorporated in the master generating routine; two-way sorting. Univac memo No. 100 (1954).
[32] Savidge, D. V., "Arranger" ten-word item 10×10, October, 1955, Remington Rand Division, Sperry Rand Corporation.
[33] EDPM program brief 3–705 sorting table; three way merge without record storage unit (times required under various conditions). IBM form 22-6644-0 (1955).
[34] EDPM program brief 8–702 sorting techniques; two-way merge. IBM form 22-6656-0 (1955).
[35] Blum, and Fogg, *Sorting on the 701*, Proceedings of the IBM Computing Seminar, May, 1954.
[36] Cline, R. L. *Sorting and Merging on Electronic Data Processing Machines*. Read Before ACM meeting (1955).
[37] Datamatic Corporation, 151 Needham St., Newton Highlands, Mass.
[38] Raycom data processing system, designed for business. Raytheon Manufacturing Company, Computer Department, Waltham 54, Mass. (now incorporated in Datamatic above).
[39] Electronics Division, National Cash Register Company, 3348 W. El Segundo Blvd., Hawthorne, Calif.
[40] Isaac, E. J., and Singleton, Richard, *Sorting by Address Calculation*. Read before ACM meeting (1955).
[41] Savasort Corporation, P.O. Box 2428, 530 24th St., W. Palm Beach, Fla.
[42] Monroe Calculating Machine Company, Monrobot Laboratory, Morris Plains, N. J.
[43] Univac Division, Sperry Rand Corporation, 1902 W. Minnihaha Ave., St. Paul 4, Minn.
[44] Friend, E. H., *Sorting*. Read before ACM meeting (1955).
[45] Fuller, H. W., Husman, P. A., Kelner, R. C., *Techniques for Increasing Storage Density of Magnetic Drum Systems*. Proceedings of the Eastern Joint Computer Conference (1954).
[46] Electronic data processing systems—functional descriptions and specification summaries. Laboratory for Electronics, Inc., 75 Pitts St., Boston 14, Mass.
[47] Soden, Walter, Sorting numbers into sequence on a digital computer, Computer Control Company, Point Mugu, Calif.
[48] The TapeDRUM new Brush rapid access high capacity magnetic memory, Clevite Research Center, 540 E. 105 St., Cleveland 8, Ohio.
[49] Hollander, G. L., *Quasi-Random Access Memory Systems*. Read before American Institute of Electrical Engineers, Columbus, Ohio (May, 1955).
[50] Rabinow, Jacob, "The Notched Disc Memory." *Electrical Engineering* (August, 1952), p. 745.
[51] "Magnetic Memory goes Juke Box." (IBM 305) *Electronics* (June, 1955), p. 10.
[52] Telecomputing Corporation, 12838 Saticoy St., N. Hollywood, Calif.
[53] Photoscopic information storage, International Telemeter Corporation R-77 (1955).
[54] Shiowitz, Marc, *Digital Computer Programs for Data Sorting Problems*. National Cash Register Company, Hawthorne, Calif. Lecture notes for summer session Wayne University, Detroit, Mich. (1954). Partly reproduced in paper read before ACM (1954).
[55] Goldenburg, D., *Time Analysis of Various Methods of Sorting Data*. M.I.T. Digital Computer Lab M 1680 (October 17, 1952).
[56] Ashenhurst, Robert, *Theory of Switching, Part I—Sorting and Arranging*. Harvard Computation Lab, Bell Labs Report #7 (January 1, 1954–May 1, 1954).
[57] Schwartz, B. L., *Criteria for Comparing Methods of Sorting*. Read at ACM Conference (September, 1955).
[58] Schwartz, B. L., *Mathematical Analysis of Sorting Procedures*—WADC Tech. Rep. 53-505, Appendix B (unclassified), available through ASTIA.
[59] Seward, Harold, *Information Sorting in the Application of Electronic Digital Computers to Business Operations*.—Master's thesis M.I.T. 1954 (DCL R-232). Available through ASTIA.

Document Processing

ROBERT H. GREGORY†

CURIOUSLY, the ACM *First Glossary of Programming Terminology* has no entry for the word "document." The *American College Dictionary* defines it as "a written or printed paper furnishing information or evidence, a legal or official paper." Without becoming lexicographers, we might observe that "write" means "to trace or form (characters, letters, words, etc.) on the surface of some material, as with a pen, pencil or other instrument or means; inscribe."

Quite clearly these definitions reflect long-existing techniques for economically reducing information to a form suitable for storage and evidential use, if necessary. The definition of a document is quite restrictive for it specifies a substance (paper), the method of making marks on it (writing or printing), and its use (furnishing information or evidence, legal or official). Actual or potential use of business documents for legal or official purposes occurs frequently. Laws and regulations covering content and retention have an important bearing on areas of document processing open to change and the rate at which changes can be made.[1] Interestingly, however, "writing" is defined broadly enough to encompass most or perhaps all of the output techniques used in conjunction with computers for preparing legible copy.

Function of Document

The traditional function of business documents is to identify the parties to and nature of a transaction and to serve as a temporary storage medium (but often for rather long periods) for input to record-keeping systems.

A small firm has a high proportion of its transactions with other firms (interfirm) so that the evidential aspects of a document for settling disputes may be fairly high. A large firm, on the other hand, has a high proportion of intrafirm transactions so that internal responsibility and managerial control are more important objectives. The record-keeping systems of individual firms, whether large or small, are usually operationally disjointed and incompatible (in view of the techniques available at any time) so that manually or mechanically generated documents serve as the connecting link for manual input for the records of both firms involved in a transaction. The notion is widely held, although some exceptions will be pointed out later, that information may be machine processable by one firm but that others must handle it manually with little or no benefits from technical developments of the past fifty years or so.

Substantial intrafirm economies have been and will continue to be achieved in data processing. A striking, but far from ultimate, development in this area is integrated data processing built around the intensive use of punched paper tape within a firm. Little has been done to date, however, to improve interfirm information processing. The originator ordinarily relies upon manual processing and almost universally the recipient is forced to do so. Important technical developments in document processing and their widespread adoption may greatly change the traditional nature and use of documents.

History of Documents

We tend to think of documents as being individual pieces of paper, either thick or thin, serving as the initial record of a transaction.[2] This has not always been the case, however, even after the widespread usage of paper in Italy and England by the 14th century. In the infancy of modern accounting some 500 years ago (and today in commercially less-advanced countries), the initial record of a transaction was made in a memorandum or waste-book. Such a book was notarized by a government official upon the user's representation that it would be kept currently although the bookkeeper was not available momentarily, and that it would contain a true and correct copy of transactions. In case of dispute, it served as *prima facie* evidence.[3] This type of record, instead of individual documents, had several important features. Paper was economized by use of variable record length, only one party to a transaction had a copy of the record, the initial steps in the record keeping were kept current and the memorandum served as the basis for someone skilled in record keeping to treat more formally, when convenient.

As another example of nondocumentary records, for a long time the English government recorded collections by making notches on split sticks. Tally sticks had some good record keeping characteristics. Both parties had unique duplicates of the transaction record (by splitting the stick and giving part to each party) which were difficult to manipulate. Pen, ink, and paper replaced tally sticks so that they were abolished by Parliament in 1782 and finally discontinued in 1826. Disposing of these

† School of Industrial Management, M.I.T., Cambridge, Mass.

[1] For example, contents, to some extent, but more particularly retention periods, for business records and supporting documents are specified by both federal and state laws dealing with income tax, insurance contributions, unemployment taxes, wages and hours, and utility regulations. Retention periods often run from one to six years.

[2] "Paper," *Encyclopaedia Britannica*, vol. 17, pp. 229–30; 1945.

[3] Paciolo, Luca, from Geijsbeek's translation, "Ancient Double Entry Bookkeeping," Denver; 1914. Reprinted in A. C. Littleton, "Accounting Evolution to 1900." New York, American Institute Publishing Co. 1933, p. 64–65.

records as "secret waste" by burning them in Parliament stoves generated a lot of heat and some light on the record disposition problem, as Bowden observes, for the building itself was burned down.[4]

Two other important, but often overlooked developments in document generation and processing in the nineteenth century were the development of carbon paper with its related duplicating techniques and invention of the typewriter. The deluge of paper resulting from these inventions is still growing.

The auditability of documents is considered an important feature whenever any change is proposed. During the thirteenth century, and for several centuries thereafter, the word "auditing" itself meant "to hear" for few could read and very few could write. Auditing consisted, in part, of a public hearing of the results of the fiscal activities of governmental officers by delegated representatives of citizens. Details of receipts were tested against common or public knowledge of what should have been collected and details of payment reports were made sufficiently public to reduce the temptation to fraud.[5] Today auditors look aghast at any suggestion that documents not be produced at certain points in the record generation process because, not too facetiously perhaps, they demand an opportunity to exercise their ability to read.

These extremely brief comments on the historical development of documents should make the points 1) that the nature of documents represents a rough economic balance of the technical factors, including the skill of people, available at any time, and 2) readability was not necessary in the early development of auditing to satisfy auditors who, in fact, might not have been able to read.

Improvements in Document Processing

An economic appraisal of the problem of improving business data processing involves balancing the costs of making a change and using the new technique against the benefits derived. Determination of the cost of document generation and processing is difficult because numerous documents may be considered imperative with only a limited amount of freedom concerning time and place of origin, nature, content, utilization, and disposition. Any particular purpose may have only by-product costs associated with it because the basic document will be developed and used without regard to the specific purpose at issue.

Assignment of value to documents is difficult because they are considered to be outside the main stream of revenue production and merely reflect what is happening. Experimentation to try to find a nearer optimal balance between document cost and value is, in itself, expensive because of the pervasive nature of documents, both within and between firms, and the cost and delay in making changes for experimental purposes.

The broad areas in which document processing is being or can be improved include the following:

1) Make the document independent of manual processing either (a) after generation or (b) at generation and thereafter.
2) Separately process the information and the document.
3) Eliminate a document by making some other document serve the purpose or by entirely dispensing with documentation.

Each of these possibilities deserves some elaboration. The first suggestion is that document can be made substantially independent of manual operations after their initial generation by means of the integrated data processing concept. Some of the facilities were already available, but a few large companies did a substantial amount of work with office equipment manufacturers to perfect the use of punched paper tape as an intercommunication and document preparation device.[6] A punched paper tape is created with the initial document so that modified versions can be made later by manual or perhaps mechanical addition of only a small amount of information (quantities, prices, etc.). The whole array of office equipment (typewriters, adding machines, calculators, bookkeeping machines, telegraphic communication devices, addressograph machines, and production recorders) are integratable so that manual operations are substantially reduced after intial generation of a document which itself may be in a reproducible form.

Integrated data processing as practiced thus far, seems to be limited to a firm including home office, outlying offices, and branches. Document processing within the firm is speeded by faster communication and reduced manual processing.

Substantial opportunities for economies lie in interfirm communication by integrated data processing. The problems of standardizing intercommunication, for example, in order to make one firm's purchase order the counterpart of another's sales order, appear to be so great that probably little progress will be made in the near future.

Important strides in mechanizing data processing are being made through the use of document reading devices. Character recognition machines are available from several manufacturers to read either one or perhaps a variety of type styles. In one verison of this technique developed for oil company credit accounting the plan is to have service stations imprint punch cards with the customer's number from a charge plate. The novel feature of the technique is to machine read the custom-

[4] B. V. Bowden, ed., "Faster Than Thought." London, Sir Isaac Pitman and Sons, Ltd., p. 246–47; 1953.

[5] A. C. Littleton, "Accounting Evolution to 1900." New York, American Institute Publishing Co., p. 262–4; 1933.

[6] H. F. Van Gorder, "Achieving greater productivity in accounting through integrated data processing," *N.A.C.A. Bull.*, vol. 35, pp. 1708–31; August, 1954.

er's number and punch it into the same card. The amount of the sale will be manually punched into the card so that the original record becomes machine processable thereafter.[7] This approach to document generation and processing is an alternative either to giving a customer a block of punched cards containing his identifying number, or to punching his number into a card at the service station by means of a punching device and a credit card in the form of a punch card master plate.[8] These three versions of solutions to the oil companies credit sales accounting problems rely upon semi-automatic generation of punch cards before distribution to customers or at the service station. Complete mechanization of document generation or its elimination by direct communication to a district office is not yet considered feasible (for example, the gasoline pump contains a small multiplying unit which might be modified to produce a suitable form output showing quantity, price, and amount) because of the cost of equipping numerous service stations and the low utilization of rather expensive communication lines.

Less sophisticated than character recognition devices are magnetic and photo electric spot readers used by several banks for travelers and personal check processing. Account identification numbers are imprinted on paper checks before release. When received at the bank on which drawn, a clerk punches the amount (for personal checks) and the reader picks up the identification number to start the processing operation. The initial operation for traveler's checks may consist of getting the number into form suitable for collating against issuance information to reconcile checks paid. The identification number and amount for personal checks may be fed into a computer for current record keeping, according to the plans announced by one bank and more fully described at yesterday's session.[9] Personal checks have to be sorted for return to the person writing them. This means that both the information and the physical evidence of the transaction must be sorted to fulfill the general specifications of the present banking system whereby documents are returned to originators.

The approach to check processing outlined above has severe limitations because only the bank on which a check is drawn can process it economically. At present, however, about two thirds of the processing task is done while a check circulates through other banks. More will be said later about interfirm data processing which is an acute problem for banks.

Department stores have a difficult problem of data processing centering upon credit sales. Every transaction requires information about customer credit and accounts receivable, inventory-sales analysis, and sales person earnings. Manual preparation of a sales slip and related operations might be eliminated by ingenious combinations of tokens or cards for each factor involved. Periodically, new techniques are announced to solve part or all of this difficult problem. Canning and others[10] described a "central records" or "delayed-processing" approach at the 1954 Western Joint Computer Conference for the department store data generation problem. More recently NCR announced that it had a solution to offer for two phases of the problem.[11]

The history of the idea of automatic data input for a centralized system for a department store is lengthy. A punch card central records system designed to solve some of these problems in essentially the same way as that proposed by Canning, considering the technology of the time, was temporarily installed in a department store about 1930.[12] Little effective work has been done since that time.

The second area suggested above for improving document processing involves separation of the information from the paper itself. This approach has large possibilities, if the input document can either be disposed of entirely or filed and replaced by a summary or detail list after processing. The opportunities are severely restricted, however, if the original document must be delivered as evidence.

Some examples will illustrate this point. United States banking practice presumes that a check will be delivered to the person paying it. Any scheme for processing the information on a check (essentially instructions for a bank to charge one account and credit another) separately from the check itself is expected to provide for sorting the checks for return to the payer. Some ingenious processing techniques have been used or proposed for making the information and paper check end up at the same point.[13] These include attaching a punched paper tape[14] or a piece of punched card to the paper checks and preparation of a duplicate punched card for every check. More recently, a committee of the American Bankers Association suggested that machines be developed to simulate manual operations in order to make check handling automatic.

Banking practices in some European countries and elsewhere does not provide for returning checks to the person writing them. If the number of referrals is relativly low, the no-return plan has some merit. Information can be taken from the check at an early stage for automatic processing by most economical techniques and the check itself disposed of as desired. Jacobs and I

[7] News Release concerning Scandex, March 2, 1955, Farrington Mfg. Co.

[8] "Sohio launches IBM credit card system," *Natl. Petroleum News*, vol. 47, pp. 81; March, 1955.

[9] "Cutting bank's paperwork," *Business Week*, p. 141; October 1, 1955.

[10] R. G. Canning and others, "Trends in Computers: Automatic Control and Data Processing." New York, A.I.E.E., 1954; "Businesss data processing: a case study," pp. 80–104.

[11] "Electronics counts the stock," *Business Week*, pp. 57–58; February 5, 1955.

[12] L. F. Woodruff, "A system of electric remote-control accounting," *Elect. Engrg.*, vol. 57, pp. 77–87; February, 1938.

[13] A survey of the proposed methods is given in R. H. Gregory, and Herbert Jacobs, Jr., "A study of the transfer of credit in relation to the banking system," M.I.T., Dynamics Analysis and Control Lab., Rept. No. 87, July 1, 1954, 26 p.

[14] G. W., Brown, and L. N. Ridenour, "The processing of information-containing documents," *Proc. West. Comp. Conf., 1953*, New York, The Institute of Radio Engineers, pp. 80–85; 1953.

made the proposal a year or so ago that the "check" (either paper, punch card or in any form suitable for transmission between paper and bank) consist of a depositor's instructions to his bank to make payments to a specified party. Upon entering the banking system, the bank's relation to a depositor could be verified and the information separated from the "check" for processing in computers within the bank and by wire circuits between banks. If a tangible check were used, it could be returned to the depositor without having traveled any further than his bank. If the instructions were intangible, the bank concerned could acknowledge the transaction without returning anything, if so desired. Our proposal, you might like to know, has evinced some interest in commercial circles but little in banking.

The Treasury Department plans to separate the information from the check itself in its payment-reconciliation procedure even though the check is a punch card. When received from Federal Reserve banks, the unsorted checks will be put through an input device to take off the check number and amount and assign the check a transaction number for any subsequent reference required. The check itself will be filed in transaction sequence and the payment-reconciliation procedure carried out in terms of the check number on a computer.

The application of management science and other techniques is supposed to improve the art of management by, among other things, reducing the volume of paper used in decision making. The 21st century manager may look at, say, a scope output of the marginal cost of producing widgets, and, if the system itself is not capable of making adjustment, the manager will do so. In the meanwhile, however, low-cost document preparation through X thousand line a minute printers will undoubtedly increase the amount of documents for busy managers to read. Relief from the flood of papers will, at best, come slowly.

The third avenue suggested above for improving document processing was the elimination of the document by making some other document serve the purpose or by dispensing with it entirely. For example, sixteen banks operate a clearing house for settlement of freight accounts between carriers. The plan provides for carriers to deposit freight bills with a bank which then transfers funds from the freight customers' to the carriers' accounts, thus eliminating use of checks and several clerical procedures. An operating improvement involves bill validation by carriers instead of shippers or receivers, thereby further reducing the procedures.[15]

The partial elimination of documents in business will result from several causes. As pointed out above, the economics of document creation and use will essentially control the situation. The cost of operating offices geared to documents in the traditional or even relatively new sense has increased greatly over recent years relative to the cost of production operations. We can look forward either to processing at lower costs or discontinuance of some documents.

Continued growth in the size of firms will reduce the need for interfirm communications but increase the need for intrafirm communications. This, however, is precisely the area where a firm has the most freedom to modify its information system including, within some limitations, the elimination of documents. Centralization of data processing at headquarters for a large firm will reduce the intrafirm communication problem because the output of one application will more often be the input of another.

The small amount of interfirm integrated data processing that has been accomplished to date indicates that it will be a long time before two firms transfer information directly from one system to another by direct interconnection. The Bureau of Standards did enough exploratory work on data processing shared by two computers to show that interconnection is feasible for nonstandardized computers, if the master computer is sufficiently flexible.[16] The possibilities for document elimination by effective interfirm communications are enormous.

Intrafirm vs Interfirm Improvements

Improvement in document processing should, in the rational world of economics, be determined by balancing 1) the cost of operating the system utilizing the new technology versus the old, 2) cost of transition from one to another, and 3) the advantages derivable from the new technology. Transition costs tend to be quite large when the technology changes abruptly—the substantial outlays by industry to study the feasibility and application of computers is an example in point. Further, important changes raise numerous uncertainties about the new system vs the old and especially operating costs, system performance, and changes in internal control relationships.

Intrafirm improvements in document processing are substantially easier to make, although I feel that the oportunities are far less striking than for interfirm improvements. Changes within the firm require the cooperation of a finite number of management and operating personnel, and auditors—both internal and external. Interfirm improvements require that these and a much larger number of outsiders (*e.g.*, customers, suppliers, creditors, etc.) affected by any change will sympathetically, at least, accept any change. It seems to be quite difficult to convince any large number of companies of the interfirm opportunities available.

The problem of processing of about ten billion checks in the U. S. each year will serve as an illustration. The precise size of paper checks has been an expensive problem in terms of extra handling, errors, low employee

[15] "Freight payment plan," *Banking*, vol. 47, p. 108; April, 1955.

[16] "First Annual Progress Report on Applications of Electronic Data Processing Techniques to Supply Management Problems," NBS Rep. 3786, Washington, D. C., Dept. of Commerce, 1954. "Data processing shared by two computers," pp. 28–33. Also M. E. Stevens, "An Experiment in Data Processing Shared by Two Computers," NBS Rep. 3602, Washington, D.C., Dept. of Commerce, 1954.

morale, poor work, and customer complaint[17] for many years. Little standardization has been achieved as the variety of sizes seems nearly as great as ever. The basic stumbling block to standardization with manual processing is that the economies available for the system as a whole are not proportionately available for individual banks. Either group action is taken or little is done. Given a sufficiently improved processing technology, however, one might expect to find standardization develop to permit exploitation of the technology. This has happened via punched cards as reflected by their fairly widespread usage. It is rare, however, for a punched card check to be processed mechanically by both the depositor and the bank. Punch card checks created by a bank for small depositors are usually manually processed by depositors, although mechanically processed by the bank on which drawn. Generally, with a few notable exceptions, large depositors create and mechanically process their punch card checks with so little uniformity that they must be manually processed as paper checks by banks. Punch card checks are manually processed almost universally in interbank operations because of the lack of standardization of card checks either issued by both large depositors of several banks or prepared by various banks for small depositors. Banks apparently consider the problems of interfirm improvement of document processing so great that they spend their energy on improving manual or simulated manual operations within their bailiwick[18] with insufficient regard to the advantages of an industry-wide attack upon problems of interfirm (banks and depositors) document processing.[19]

Precision in Document Processing

A number of companies report economies in document generation and processing by reducing the precision of the data involved.

Statistical quality control techniques enable determination of the quality of data generation, and if out of control, furnish clues to find the source of difficulty at far less cost than complete duplication or 100 per cent inspection. Bear in mind, however, that quite different problems are involved in determining the level of quality and developing a desired quality. Several companies, including Standard Register, Aldens Inc., and United Air Lines use statistical quality control techniques to achieve a better balance between the cost of document preparation and the value of the resulting quality.[20]

Study will probably indicate that identical processing is not required for all documents of a given type so that strata may be developed for different processing methods based on, say, dollar amounts involved. For example, one bank postpones the inspection of signatures on checks for less than $500 until after the checks are paid.[21] The American Hard Rubber Co. reported that it did not verify the accuracy of its invoices for less than $1,000.[22] The controller feared that his punch card preparation of invoices might result in fantastic amounts; not that invoices would be too small. A small error ratio experienced by one company warranted adoption of a plan for checking carbons of customers bills after originals were mailed in order to avoid delay. Few mistakes arose, but the potential damage was large so that customers were advised of mistakes.[23]

An error analysis that I made several years ago indicated that invoices received by one firm were probably not worth verifying if less than $500. In fact, invoices for small amounts, as prepared by vendors tend to have a downward bias so that determination of errors often increased the amount to be paid.[24]

Reduced precision in document preparation and record keeping is probably suitable for numerous business operations. Rarely are decisions made on the basis of four significant digits and often less are available. Reductions in precision apparently have not gone beyond eliminating pennies and then almost exclusively in internal records. External transactions are usually continued with penny precision. Cents elimination is a step in the right direction, but might be extended to still more digits and applied to interfirm transactions. The idea was adopted by business only recently and has not spread widely because the cost of rounding a number seems to be too high with the office machines and methods in use. Reduced cost of rounding by computer programming and the higher cost of storage may result in wider adoption.

In summary, document processing techniques in use at any time reflect an economic balance of the factors available. New developments are adopted after a time lag that is increased by the cost of introducing the new technique including the cost of disruption in changing from one to another. Change is accelerated by large relative gains (considering both costs and revenues) from the new system in relation to the old. Realistically, nearly all techniques of document processing, from quite old to very advanced, will continue in use by business at large and even within a firm for different purposes.

[17] R. C. Deering, "Co-existence with automation," *Banking*, vol. 48, p. 58; September, 1955.

[18] Another development, and probably a noteworthy one, in bank processing of check is ERMA developed by Stanford Res. Inst. for Bank of America and discussed at length elsewhere in these proceedings. Brief details are given in "Cutting bank's paperwork," *Business Week*, p. 141; October 1, 1955.

[19] Bank Management Commission, Savings and Mortgage Division, "Automation of Bank Operating Procedures," New York, American Bankers Association 1955.

[20] "Controlling clerical errors," *American Business*, vol. 24, p. 13; February, 1954.

[21] A. O. Dahlstrom, "Using numbers instead of names for depositors," *Banking*, vol. 47, pp. 48–49; August, 1954.

[22] J. A. Sonnischen, "Electronics in the Office Problems and Prospects," Office Management Series, No, 131, New York, American Management Assn., 1955; "Preparing for electronization: what the smaller office can do now, part I, streamlining operations," pp. 14–22.

[23] Phil Hirsch, "Job formerly requiring 300 hours now completed in 50 hours," *American Business*, vol. 34, p. 19; June, 1954.

[24] R. H. Gregory, "The accuracy of invoices received for payment," *The Controller*, vol. 21, pp. 11–13; January, 1953.

Original Documents in Retail Accounts Receivable

V. H. ROMAN†

IT IS indeed a pleasure to discuss the retail accounts receivable function. This is especially true as the Eastern Joint Computer Conference has not generally included the so-called paper processing type of subject in their program and we who have a prime interest in this particular field are always happy to spread our story as widely as possible, especially among individuals such as you who are primarily interested in the application of new techniques to all types of business problems.

Before getting on with the subject let me briefly outline the part economics plays in the picture. In the so-called data processing or paper handling area of activity which includes retail accounts receivable and other clerical and accounting work generally, the economic benefit resulting from introduction of new techniques is one of the major factors of justification. Unlike the research, technical, or management decision fields of operation, clerical and accounting activities usually involve the application of considerable manpower to handling a large volume of paper and any improvements, methods-wise, are primarily in the direction of reducing this manpower cost. Any cost reduction is consistent, of course, with providing for management the information it requires sufficiently current, to operate the business most advantageously.

The design and operation of a retail accounts receivable system will differ with its application to the various kinds of industry. The petroleum industry credit card activity has been selected as the subject for this discussion. It was a good choice as it involves a sizable volume activity, the detail of which generally originates among a large number of retail outlets located over a wide geographical area. In the larger companies the credit card operation represents their largest paper processing job and requires more clerical help than any other function. It also follows that accuracy and a desirable presentation of the billing is a must from a customer relations viewpoint.

Principal requirements of an oil company accounts receivable system include:

1) Issuing the customer a form of authorization to obtain products on a credit basis. This may be a credit card, with an account number of 7 to 10 digits, issued quarterly, annually, or for some other specific period. Another form is a book of invoices preprinted to indicate the customer's name and credit card number or prepunched with tabulating equipment to indicate the credit card number and other significant data.
2) The accurate and intelligible preparation of the charge invoice at the service station in the minimum amount of time.
3) Communication or physical transfer of the completed invoices to the accounting office.
4) Processing of the charges in the accounting office, handling customer cash remittances and preparation of monthly billing for mailing to customers with copies of the invoices representing the current month's charges.
5) Preparation in the accounting office of various reports for credit control and statistical requirements.

For the purpose of discussing the subject in some detail let us visualize a field accounting office which has 150,000 active credit card accounts and processes 900,000 invoices and 150,000 customer remittances per month. Billing is on a 20-cycle-per-month schedule. On a daily work load basis this volume represents the processing of 45,000 invoices and preparation of 7,500 customer statements.

Under a fully manual system the customer is supplied with an addressographed or typed credit identification card. The service station salesman is required to prepare an invoice for each delivery writing the customer's name, credit card number, date, and the detail of products purchased with price per gallon and total amount of the purchase.

Upon receipt of the invoices in the accounting office they are comptometer-added to reconcile with the invoice transmittal report, then processed and manually sorted for maintenance of cycle controls and to arrange them in account number sequence. Sorted invoices are then filed behind respective customer account guide cards pending cycle closing and customer billing.

Billing is accomplished by comptometer or adding machine addition of the charges for each account and entering the total on a statement addressographed with the customer's name, address, and other indicative information. The statement with copies of current month's charge invoices is mailed to the customer. The file record may be a duplicate copy of the invoices or a microfilm record of the originals. In most systems a copy of the customer statement is used as the accounts receivable ledger. Another copy of the statement is used for credit administrative purposes and is manually processed to indicate the age status of past due balances.

Customer remittances are handled currently and at each billing cycle closing any unpaid balances are carried forward to the new statements to indicate the previous balance owing.

† Manager, Organization and Methods Div., Standard Oil Co. of Calif.

This brief outline of a manual procedure does not do justice to the job in keeping with the manpower requirements. An approximate force of 100 people is required to perform the clerical and accounting detail associated with 150,000 active accounts and the processing of 900,000 invoices and 150,000 remittances each month.

The foregoing, while related to a manual system, does of course include the use of addressograph equipment for imprinting statements and customer credit cards; but it does not provide any form of mechanization at the service station or in the accounting office for automatic handling of the invoice paper volume, statement preparation, and related detail.

In any area where the volume of documents to be processed reaches the magnitude we are discussing there exists a more or less constant search for equipment and methods which will more economically provide the necessary end results. Some progress has been made and recent developments in the electronic equipment field give much encouragement for the future.

Before commenting on the more recent developments I would like to briefly outline some punch card techniques which are now being used. One method is to provide each customer with a book of tabulating card invoices prepunched with the customer account number and preprinted with the customer name. They are presented at the service station where the station identification, date, and detail of purchase is entered on the card. The 51-column portion of the card invoice is given to the service station attendant and the stub portion is the customer's record of purchases made. Upon receipt of the invoices in the accounting office the amount of the purchase is key punched into them and they are ready for machine processing through the various steps with conventional punch card tabulating equipment. This system has both good and objectionable features when viewed from the customer relation, accounting, cost, and other viewpoints to be considered.

Another system recently introduced provides the customer with a plastic credit card in which the account number is punched. Each service station is equipped with an imprinting and punching machine into which the credit card and tabulating card invoice are inserted. The machine senses the account number from the punching in the credit card and mechanically punches it into the tabulating card invoice along with imprinting other identifying information. Upon receipt in the accounting office processing is similar to the system previously mentioned.

Still another approach now in the process of development is a credit card in which the account number is punched, and supplying the service station with a machine which may be set to punch both the account number and the amount in the tab card invoice. Processing in the accounting office is as previously outlined.

One other system used to considerable extent with a manual or tabulating accounting procedure is the issuance of a credit card with a small printing plate, indicating the customer's name and account number attached to the reverse side of the card. The service station is provided with an imprinter which imprints the invoice with the customer's name, account number, date, and other identifying information from the credit card. This approach is intended primarily to reduce the time of the customer at the station. It does, however, also materially aid in the accounting office processing by providing accurate and easily read account numbers for manual sorting or punching purposes.

Any system which incorporates the use of a machine at the service station must give full consideration to the type and cost of this equipment. The type of machine is important from a maintenance and registration tolerance viewpoint. It must be durable and simple to operate, and provide against registration differences beyond tolerances acceptable in subsequent accounting office processing. The initial cost of the machine must of course be within the range of economic justification with consideration to the large number required for equipping the thousands of retail outlets.

More recent developments toward complete automation in the retail accounts receivable field include the use of electronic equipment, such as character sensing machines, high speed sorters and line printers, and key punch equipment having special facilities to expedite the punching requirements. Now let me briefly describe a selection of equipment to perform the accounts receivable operation utilizing these new developments to achieve maximum automation results. The machine selection is predicated upon using specific equipment for specific purposes in keeping with the maximum input and output speeds and weighing the cost of performing the over-all job.

The credit authorization is a plastic card embossed to indicate the customer's name and account number. The last digit of the account number is designed as a check digit and represents the unit number of the sum resulting from a mathematical formula applied to the preceding digits of the account number. The service station is equipped with an imprinter which is used to imprint the embossed customer name and credit card number from the credit card on a 51-column tabulating card invoice. A paper carbon copy of the invoice is given to the customer at time of delivery.

When received in the accounting office the imprinted tab card invoices are first processed through the electronic character sensing equipment which reads the account number, checks the reading with the check digit of the number and, if in agreement, actuates the automatic punching of the number in the invoice. The speed of the electronic sensing and checking is in excess of 200 cards per minute and is presently much beyond the speed of available punching equipment.

In those instances when an imprinter is not available and the customer account number has been manually entered on the invoice it is manually punched in the card with a high speed key punch machine equipped with a device which verifies the number with its check digit, thus eliminating the necessity for otherwise verifying the accuracy of the punching or the number itself.

For punching the amount of the charge in the card invoice a key punch with adding machine and printed tape attachment is under development. The tape with subtotal and total printing will allow for balancing the invoices with the transmittal reports received from each service station, and simultaneously serve to verify the key punched amount.

All sorting to establish cycle segregation and for arranging the invoices in account sequence can be accomplished with a high speed Electronic Sorter operating at a speed of 1,000 cards per minute.

For addition of the invoices and cash remittance cards to establish cycle and other type controls an Electronic Statistical Machine with an input and output speed of 450 cards per minute is available.

Preparatory to the billing operation the work of carrying forward previous unpaid balances, application of current month's cash, summarizing of current month's purchases and the preparation of a current month's balance card with past due amounts properly aged may be performed with an Accumulating Reproducer having an input and output speed of 200 and 100 cards per minute respectively.

The customer statement is a tabulating card with the 51-column portion designed to be mailed by the customer with his remittance to the accounting office. Preparation of the statement is performed with a Line Printer equipped with a bill feed device for feeding the card form statements. The speed of existing conventional tabulating card line printers leaves much to be desired. One hundred or 150 lines per minute does not keep pace with other developments in the equipment field. However, a new 900 line per minute printer has been announced, first deliveries to be made in early 1956. The design of this equipment with its summary punching facilities will meet the requirements of the accounts receivable function and provide the input and output volume required. With such a machine the step just referred to in which the previous month's balance and current month's charges are summarized and the past due balance is aged on a new summary card could be discontinued and the summary operation combined as a by-product of statement preparation.

The methods just described represent the combination which will, in my opinion, do the particular job most economically and meet customer and management requirements. While some of the equipment, such as the electronic character sensing machine, new type key punches, and high speed electronic line printer, have yet to be tested under the users' operating conditions it appears all of this equipment will prove satisfactory and be available in the near future. Some of the other machines referred to and the methods outlined are already being used in some of the punch card mechanized applications throughout the country.

As stated at the outset of this discussion, the type of accounts receivable system will differ with its application to various industries and also with different kinds of accounts receivable problems within an industry. In the case of credit card accounting, we are faced with a very large input problem, very simple computing requirements, and a sizable output requirement. The need for interim data on specific accounts between billing periods is practically nonexistent and there are no sales analysis, inventory control, or spot reference requirements involved. We are faced primarily with a very large data processing chore of sorting and adding invoices, application of cash, and preparation of monthly statements and credit data.

The possibility of utilizing a large scale electronic data processing system for this work has been considered and will warrant continued close follow-up and analysis. It appears to me that a machine designed especially to meet the requirements must be conceived in order to keep the cost factor at a level where the economics will favor such equipment.

Discussion

J. B. Peistrup (Shell Oil): Will you give estimated savings in key punch operations through utilization of electronic scanner?

Mr. Roman: Whether statistics will hold up I don't know but we figure another ten per cent reduction of over-all manpower as a result of scandex.

E. G. Aghib (Remington Rand): Could the speaker furnish details and the manufacturer's name of the plastic plate charge card punching device?

Mr. Roman: Farrington Manufacturing Company, right here in Boston. The embossing work itself can be accomplished with an addressograph graphotype machine having a special set of dies in it. Of course, where we intend to use scandex the dies have to be uniform for scanning purposes.

J. Svigals (IBM): How much does it cost, per account or per transaction, to process data under the present system?

Mr. Roman: One clerk is required for approximately 1,400 accounts. On this basis the cost would be the rate of pay applicable to a file clerk type of activity.

J. Thomson (RCA): Do you foresee the time when it will be unnecessary to provide the customer with a copy of each transaction document, keeping these on file in some form for reference as required?

Mr. Roman: I wish I could be more optimistic on that score. We have some situations that differ from other types of accounts receivable systems. Department stores, being more centralized, normally include all charges on the statement for the month in which purchases were made. Bank commercial account statements are quite similar. Credit card accounts receivable are quite different. Our cards are honored by selected companies in other geographical areas and *vice versa.* Credit deliveries made while on an extended trip do not all reach the customers' billing office in sufficient time to be included in the statement for the month in which purchase was made. To give the customer a statement indicating the total monthly billing would not be sufficient, without copies of the invoices, for reconciling with his copies of invoices received at time of delivery. For that reason, while we would like to get away from returning the tickets, I don't foresee any relief in this area.

G. E. Gourrich (Telecomputing Corp.): How are these efforts being coordinated so that the information produced at the station can be fed into the electronic computer when many people work on the system?

Mr. Roman: Various companies are developing systems to meet their respective problems.

Mr. Gourrich: With the manufacturers.

Mr. Roman: The manufacturer has not yet been brought into the picture specifically. Equipment we have looked at to date is not economically justifiable. Anything large enough, memory-wise, is too costly.

Mr. Gourrich: Could you give any guide as to cost requirements for a point of sale unit at the gas station?

Mr. Roman: The cost requirement for a punching unit at the gas station varies anywhere from $30 to $60, and there is one under development, to cost approximately $90, which will also punch the amount in the card.

The Computer and Its Peripheral Equipment

NATHANIEL ROCHESTER†

Introduction

MULTIPLICATION SPEED was once an accurate measure of the speed of an automatic calculator, because the prominent problems of that earlier day required a lot of multiplication but had only a little data, demanded relatively little memory, and had short answers. Today, things are profoundly different. One now speaks of problems limited by tape, still others limited by printing, logic, and so forth. Behind these statements is a tacit understanding that some of these limits would recede if high speed memory were larger. This situation is discussed in this paper with particular emphasis on the peripheral equipment which surrounds the largest currently available machines.

Beginning with problems which are mostly calculation, the discussion proceeds by successive stages to problems which are almost entirely input and output. At each stage there is a description of relevant equipment that is about to become available or that has recently become available. Most of the examples are IBM machines because the author had access to current information about these and could thereby make the illustrations informative and precise. In conclusion there is a treatment of a number of special topics.

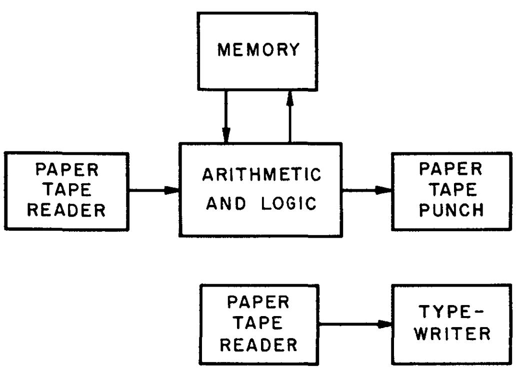


Fig. 1—Typical early calculator.

Partial Differential Equation

Fig. 1 shows a typical calculator of a few years ago. This example establishes a reference point from which to start, although it differs from subsequent examples in that it does not represent a powerful machine that would be widely useful today. A problem, which this machine could do, was a partial differential equation such that both the program and the array of numbers could fit in memory at once. The operation consisted of loading the program and data into memory, performing a few dozens or hundreds of iterations in each of which the data were worked over to make them fit a little better, and finally the resulting data were punched out on tape. Later the tape was run into a separate typewriter and the answer printed.

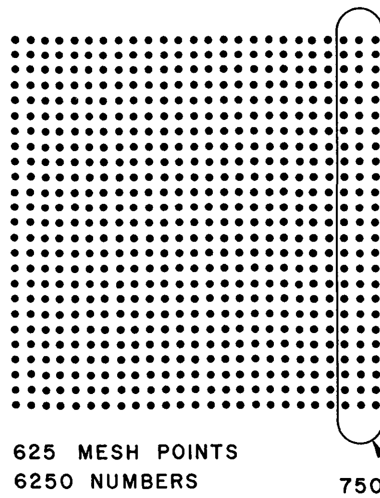


Fig. 2—Partial differential equation.

Today a large partial differential equation just will not fit into present high speed memories. The example shown in Fig. 2 is a hydrodynamics problem. There are 6,250 numbers which have to be worked over. The calculator must go over the mesh about 1,000 times and, in each trip over the mesh, stop at each point to make a 30 millisecond calculation. The mesh could be stored on tape or drum; tape has been chosen for this illustration.

The method is to read the first 750 numbers into the calculator and make one set of calculations. Then the first column of 250 numbers is written on tape and the next column read in. This goes on until the whole mesh has been covered to complete a single iteration. 1,000 such iterations might be needed. The balance of time is shown in Table I.

TABLE I

Activity	Hours
Calculate	4.4
Read and write magnetic tape	.4
Print one line every iteration	.1
Print 125 lines every 10 iterations	1.6
	6.5

Most of the time is spent in calculation, as it should be for this kind of a problem. The next largest amount

† Manager of Information Research, Engineering Laboratories, International Business Machines Corp., New York, N. Y.

of time is devoted to printing out the whole array of numbers periodically. On the IBM 704 this can either be done on a directly connected printer or written on magnetic tape and printed later on a separate tape operated printer.

The magnetic tape consumes only 6 per cent of the time; however paper tape could not have been used because it would have taken 150 hours or more. The present trend is to go to larger fast memories and eliminate tape from problems like this. The purpose is not so much to eliminate the 6 per cent of time lost running tape, but to make possible some other methods of problem solution. The limitation, imposed by tape, of having to go through the mesh in an orderly fashion prohibits some mathematical techniques. In keeping with this trend, a 32,768-word core memory is being designed for the 704.

Inventory Control

Many scientific problems require more running of tape and less calculation. However, it is most convenient, at this point, to go over to commercial problems. Ordinarily, commercial problems require more input and output while scientific problems require more calculation. Consider the inventory control problem shown in Fig. 3.

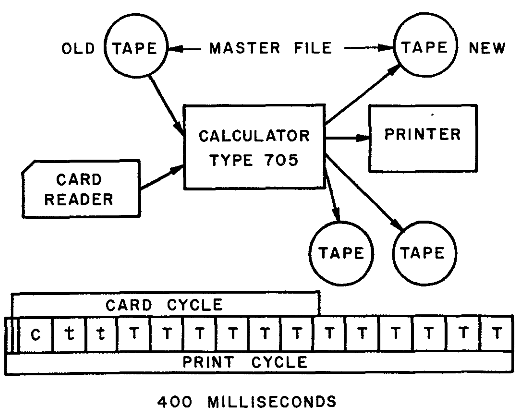


Fig. 3—Inventory control.

At the start of the operation the main inventory is on the old master file tape. During processing, data are read from this tape, updated, and written on the new master file tape.

The punched card reader is used to introduce transactions. For each transaction, such as a withdrawal from stock, there is a punched card. These cards are mechanically sorted to the same sequence as the main file before the file maintenance operation begins. Ordinarily only some of the stock items are active. In this example a rather difficult condition has been chosen because only one item in ten is active.

The sequence of steps begins with the reading of the first transaction card. Then the first master record is read. If the part number does not match that on the transaction card, this record is written on the output tape and the next record is read by the calculator. The appropriate IBM data processing machine for this operation is the Type 705 which would do this reading and writing simultaneously. When the correct master record was found, it would be updated with the information from the card. It is assumed, for this example, that on this occasion, three reports on the activity are prepared simultaneously. One would be printed directly, and the others recorded on tape to be printed later. When the active record was found, records would be written on most or all of these outputs as well as on the new master file.

The utilization of time is shown in the lower part of Fig. 3. This timing chart is drawn for a medium speed printer which prints 150 lines per minute. A single print cycle takes 400 milliseconds. Simultaneously, the card reader runs, and since it is faster, it spends part of its time waiting for the printer. During each print cycle the central processing unit must communicate with the card reader and printer. The time required for this is indicated at the left of the middle timing bar, but the 5.5 milliseconds required is too short to show up well. Then there is some time, indicated by *C*, for calculation, and time, indicated by *t*, to write two output records on tape. Time is also required for a variable number of master records. As indicated by the spaces marked *T*, if no more than 12 are needed to reach the next master record which matches a transaction, the printer will not be delayed. If there are 12 or more inactive records in succession, the printer would have to wait because the problem would be tape limited.

This is an appropriate point to discuss the question of whether the card reader and printer should be directly attached to the calculator or whether the calculator run should be preceded by a card to tape operation and followed by a tape to printer operation. It is possible to disconnect a card reader with both its control unit and a tape unit from the central processing unit and operate these separately as a card to tape machine. The central processing unit can then work simultaneously on another job. The same thing can be done with the printer. If this is done these various units can all operate at once and no one has to stop and wait for the others. Often, but not always, this will reduce the cost of such a job. Usually it will increase the elapsed time from the reading of the first card to the printing of the last line. Also, it will require more human intervention in the form of moving reels of tape around and scheduling three instead of one machine operation. The choice, then, often depends upon whether one considers speed and convenience or cost to be the controlling factor. At other times no compromise is needed because one mode of operation or the other will be superior in every way.

While there are a great variety of forms of problems, the form illustrated on Fig. 3 is one of the most common. There is a master file containing business records to be read and written; and a detail file containing current transactions to be read. Also, there are almost always some outputs of current transactions. These are things like invoices, orders, paychecks, or premium notices, reports to inform executives so that they can run the business, and data for subsequent machine operations. This basic form has many minor variations. There may be several detail files to be read and merged as the operation proceeds. There may be a variety of current outputs. There may even be more than one master file, although with the advent of magnetic tape this has become less common. Since the subject is mainly machines rather than problems, this common form of problem will be used to illustrate the next stage of increased emphasis on input and output.

Premium Notice Writing by Multiprogramming

Some parts of premium notice writing, in insurance, are like the previous example but there are apt to be about 100 inactive records in the master file for every active record. Notice that even in the inventory problem the central processing unit was spending most of its time running tape and only a little time calculating. This next stage involves a problem which would be ten times worse. It is clear that different equipment is in order.

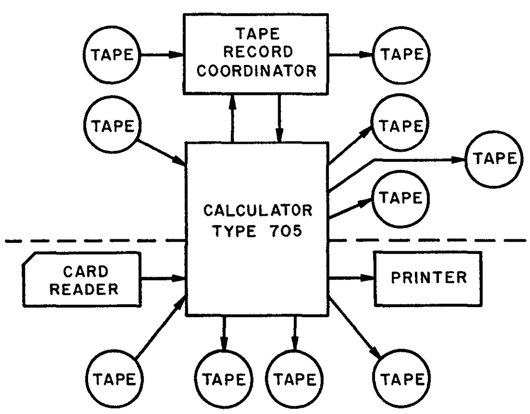


Fig. 4—Multiprogramming.

Fig. 4 illustrates a system for efficiently handling even this extenuating case. The trick used is called "multiprogramming" and it involves doing two or more different jobs at once. The peripheral equipment above the dotted line in Fig. 4 is involved in the difficult file maintenance job while that below the line is involved in a different task. The calculator in the middle takes turns helping the two different teams of peripheral equipment. The Tape Record Coordinator at the top runs its master tape units almost autonomously while passing over inactive records. Then when an active record is found, the calculator briefly interrupts the work it was doing below the line to deal with this active record. Multiprogramming allows the calculator to spend most of its time on the job below the line (Fig. 4) and also allows the master file at the top to run almost continuously.

Now consider this in more detail. The Tape Record Coordinator has a 1,024-character core memory. It can do quite a few different things but it is not relevant to describe them all here. What it does in this application is to simultaneously read from one tape and write on another. It does this in response to an instruction given by the calculator. When it has finished doing this, it signals the calculator. The calculator picks up the first few characters from the record in the core memory of the Tape Record Coordinator. These should be arranged to contain the indicative information so that the calculator can see if it has found the record it wants. If so, the calculator transfers the whole record to its own memory, does the necessary calculation, returns it to the Tape Record Coordinator, sets the Tape Record Coordinator going again, and returns to the other job. If not, it merely sets the Tape Record Coordinator going and likewise returns to the other job. When it is passing over inactive 250 word records, the calculator needs to spend less than $\frac{1}{2}$ per cent of its time supervising the search. All the rest it can devote to the other job.

Multiprogramming often requires that both programs be in memory at once. This requirement is one of the factors that led to the optional 40,000-character core memory of the Type 705 calculator.

Random Immediate Access

The extreme case of a poor ratio of active to inactive records comes with a large file, say one million records, when access to one record is needed. One of the more awkward versions of this problem arises when a customer is waiting on the telephone and a file reference is needed before finishing the phone call. Sometimes it is even necessary to bring the file up to date for that transaction before answering the telephone again. The next phone call might be from that same customer's wife, and it might be necessary to know what the man had done in order to converse intelligently with his wife.

An older solution to this problem is to have a large card file and have a clerk select the appropriate card while the customer waits. If soon after the transaction, another clerk attempts to select this same card from the file he will find that it is missing. Thus he will learn that there has been a recent transaction.

With a tape system there are expedients to alleviate this problem. Books can be printed to describe the file as of a certain date, and addenda can be published to keep it up to date. However, the proper solution is to have a large enough random access file memory. A step in this direction is IBM's 5,000,000-character Random Access Memory which has an access time that varies between 0.2 and 0.6 sec. This can be regarded as a harbinger of bigger things to come.

This last example completes the sequence of problems which cover the gamut of requirements from problems which are primarily calculating to problems which are primarily data flow. The treatment has, however, by-passed a relevant matter of great importance.

Flexibility of Peripheral Equipment

Peripheral equipment can be divided into two classes. The first class, discussed above, consists of equipments which communicate by wire with the central processing unit. The other class, to be discussed now, is auxiliary units. There is a fundamental principle to be expounded, but first some examples are needed.

One typical requirement for auxiliary units arises in payroll. One usually does not want to print any checks until the payroll calculation is complete and the controls and checks have balanced. However, once the calculation is complete, it would be very wasteful to tie up the calculator while printing checks. In the 705 system this necessity can be met by disconnecting the printer, its control unit, and one tape unit from the central processing unit and connecting them as a separate installation as shown in Fig. 5. This printer has a speed of 150 lines per minute.

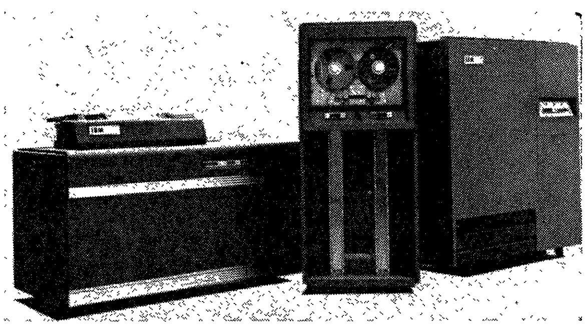


Fig. 5—Medium speed tape to printer.

When a faster printer is needed there is a 500 line per minute printer and two different 1,000 line per minute printers. Fig. 6 shows such an installation. There are two tape units, a storage and control unit, and a printer. Notice the two tape units. Only one runs at a time, but because there are two, printing need not be interrupted while changing tapes.

The tape units may be disconnected and the high speed printer with its control unit used as direct output on a calculator. Alternatively the printer may be disconnected and the storage and control unit used as a buffer between the tapes and the calculator. The advantage of this connection is that information may be transferred between the calculator and the storage unit at the rate of 43,500 characters per second whereas the tape speed is 15,000 characters per second. Therefore, the use of a tape buffer allows simultaneous calculation and running of tapes.

For a larger, more powerful installation, the Tape Record Coordinator mentioned previously may be used as a buffer. The rate of information transfer between it and the calculator is 111,000 characters per second, thus reducing the time used by the calculator for tape reading by a factor of 7.4.

The point illustrated by these various connections of peripheral equipment is that there are usually several different requirements to be met by a data processing installation. In order to meet them most efficiently, multipurpose units are needed. The examples illustrate a few of the many possible ways of interconnecting these devices. An installation can not only be tailored to the requirements of a particular location but can also be reconnected from hour to hour to meet the varying needs at that location.

Fig. 6—High speed tape to printer.

The exponents of special purpose equipment often overlook the fact that at nearly every installation there are a variety of jobs to be done. If special purpose equipments are used, they must be small enough to run continuously on each of the special purposes and there must be enough different equipments to handle the different jobs. Even this is not usually enough because the work to be done at the end of the year may be very different from the work to be done at the end of the month, which in turn may be different from the work in the middle of the month. The stored program solved this flexibility problem for the central computer; an important current trend is toward peripheral equipment with corresponding flexibility.

Another consideration is maintenance. If the peripheral equipment can function either connected to the central processing unit or separately, it is possible to service the input-output equipment of a calculator without tying up the calculator.

Grouping

No comprehensive treatment of current trends in peripheral equipment could omit the subject of grouping records on tape. It is necessary to leave blank spaces on tape between sections filled with information to provide a place for the reading head when the tape stops.

They are necessary because the signal that comes from the reading head is proportional to the tape speed. Therefore, the signal goes to zero when the tape stops. If there is recorded information under the head when the tape stops, that information is lost. The blank spaces, $\frac{3}{4}''$ long, are sufficient to record 150 characters. Therefore, if the system is to realize a major fraction of the potential tape speed, the sections of recorded information must be long compared to 150 characters.

Fig. 7 shows the efficiency of tape as a function of the length of recorded information. If the length of recorded information is 150 characters it will equal the space between records. This will mean that half the time devoted to running tape will be lost and the efficiency will be 50 per cent. Using the full 1,024-character capacity of the Tape Record Coordinator as the record length yields an efficiency of 87 per cent.

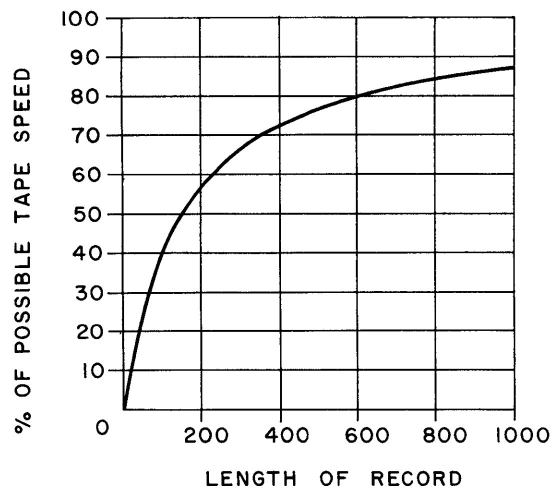


Fig. 7—Efficiency of tape.

The requirement of long records for tape efficiency runs counter to a requirement for calculator efficiency. For efficient data processing and for conceptual simplicity planning and programming the job, one wants the record length tailored to the application. Few applications have record lengths that are long enough to realize most of the potential tape speed. Therefore, provision has been made to handle groups of records. From the point of view of the tape unit grouping seems to provide long efficient records. However, the programmer and the central processing unit retain the advantages of flexible record length.

Provision has been made in several places in the Type 705 and its peripheral equipment for grouping. The notation on magnetic tape is shown in Fig. 8. Here an ungrouped tape is shown on the top and a grouped tape below. Within a group the records are separated by special symbols called record marks, which are copied into memory when the tape is read. The groups are marked off on tape by blank spaces. In memory a special character is used as a group mark. Certain instructions have features which make it convenient to deal with groups inside the calculator. The Tape Record Coordinator and other peripheral equipment also have features to simplify grouping. When short records are appropriate it is usually best to group records.

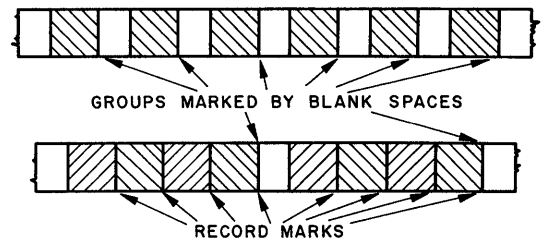


Fig. 8—Ungrouped and grouped records.

Conclusion

Current trends in peripheral equipment for automatic calculators can be summarized as follows:

1) The communication link between the machine and its environment is being strengthened so that machines can automatically handle a wider range of tasks.
2) Flexible peripheral equipment can now function as an integral part of a large calculator and then be disconnected to function separately while the calculator does other things. This provides a means of adapting a data processing installation to the wide range of jobs encountered at a typical single location.
3) File reference jobs with a very low ratio of active to inactive records have become practical with the advent of multiprogramming. This method utilizes special peripheral equipment which handles the time consuming parts of file reference while the calculator busies itself on other work. The full power of the calculator is applied to the short, hard parts of a file reference job but is not wasted during the long easy parts.
4) Effective magnetic tape speed has been rising, not only by increasing intrinsic speed, but by buffering and grouping, both of which increase the percentage of the intrinsic speed that is realized in practice.
5) Peripheral random access memories with millions of characters of data and with access time of a fraction of a second are becoming available and a trend toward even larger peripheral memories is evident.
6) A counter trend toward larger high speed internal memories is taking some kinds of work away from the peripheral equipment.

The impact of these trends is to allow automatic calculators to shoulder an increasing burden, to relieve people of more and more menial tasks, and to steadily raise our standard of living.

Discussion

L. W. Armstrong (Dept. of the Army): Please discuss the place of an auxiliary file maintenance sorter-collator machine in a large scale system; for example, selection of active records from master file for your stock control problem?

Mr. Rochester: I think that question refers to another solution of the problem which is handled by the Tape Record Coordinator. On a file maintenace job where there is a small percentage of active records you can't allow the calculator to wait while reading inactive records; so one solution is to have a cheaper device which runs through the main file and selects active records. These can be run into the calculator and it can be run efficiently.

E. L. Harder (Westinghouse Electric Corp.): Some special purpose computer applications are dictated now by waiting time to the main computer. What break-in schemes are being developed?

Mr. Rochester: A very nice arrangement, of course, would be to have some scheme so that the action in the calculator is automatically interrupted whenever something interesting is found. But a problem arises in that you can't necessarily interrupt what is going on in the calculator. If there is something in the accumulator you can't break in at an arbitrary time.

For that reason we chose one of the several solutions. That is to have the calculator periodically ask the peripheral equipment, "Are you ready?" This doesn't seem to waste very much time.

K. Cutler (Liberty Mutual): You mentioned that it is possible to have five million characters of random access storage. I do not see how this can help a company with large master file volume. Does IBM plan on increasing this capacity? Otherwise the cost would be tremendous.

Mr. Rochester: That is perfectly correct. The five million character memory doesn't help somebody with a much larger file. However, we are going by easy stages into this problem but I can assure you that eventually the computer industry will provide you with very, very large random access memories which will be large enough for your biggest files. What I can't tell you is just when this will happen. I hope that the five million characters we have will be an indication of good faith on our part.

R. S. Gillett (National Security Agency): Would not large capacity magnetic drums with direct loading from tape provide a satisfactory intermediate storage with random access?

Mr. Rochester: Yes, large magnetic drums are one of the possible solutions. What these various solutions amount to is that you have magnetic material and you have information stored in this magnetic material, and somehow you have it arranged so it moves mechanically by the head and there are quite a number of schemes for this. You can imagine magnetic plates, or rods, or discs. Our choice in this was discs but other people have chosen other things. I would say they are all roughly comparable; although you can see that we figure the discs are the best.

E. M. Stone (No affiliation): The percentage of time, effort, money, and emphasis devoted to peripheral equipment, rather than to the computer itself, seems to be increasing. Can you give some idea when this shift in emphasis started and to what extent it has gone?

Mr. Rochester: Well, I think this shift began right at the beginning. In other words, first it was realized that it was possible to make the machines which could do anything that you were willing to instruct them to do within some very wide limits. The early machines had the characteristics of being universal machines. People thought that was enough of a hurdle. I think the very first machines in 1946 had the disadvantage that you couldn't really realize the potential of the internal machine because you couldn't give it information in and out rapidly enough. So the computer industry has been hard at work on this ever since. Probably right now half of a large computer is peripheral equipment and the other half is the central machine. I think this is a reasonable ratio.

G. Wence (Social Security Administration): A tape file of 120 million accounts is to be updated once each year. 10,000 references per day must be made to the file. Is not special purpose equipment indicated?

Mr. Rochester: One problem I discussed had one active record in hundreds. This problem has one active record in 12,000. There is no question that you wouldn't want to tie up a big calculator waiting for all these inactive records to file by. I am sure special equipment is indicated there. I am sure once you found the one record in 12,000, you would want to bring the full power of the calculator to bear. Therefore it is nice to have the other equipment to help in this long searching time.

Computers With Remote Data Input

E. L. FITZGERALD†

Introduction

THIS DISCUSSION concerns the current status of remote input technology for commercial data processing installations. It deals with a few of the problems and implications of 1) the electro-mechanical recording of data on a storage media at a remote geographical location, 2) the transmission of the recorded data to a central computer, and 3) the conversion of the transmitted information to a form suitable for computer input. This general purpose approach precludes consideration of special purpose, on-line computer applications.

† Graduate School of Industrial Management, M.I.T. Cambridge, Mass.

The balance between allowable cost and acceptable input data time lag determines the techniques used to accomplish remote input. When the original recording media is acceptable as computer input, punched cards or magnetic tape for example, and the time value of the data is low, mail or messenger service can be used effectively. The only significant difference between this technique and central document recording is the induced time lag. As time requirements, either real or imagined, become critical, other means of data transmission must be selected. Cost per character of input jumps sharply as teletype, voice frequency telephone circuits, radio, or microwave channels are designed into the data flow system. This additional cost may be compounded by a lack of flexibility in the choice of characters or symbols transmitted and the necessity for conversion equipment at the terminal computer. After the initial step to higher speed transmission, additional increases in speed and allowable cost are accommodated by switching between media or adding parallel transmission channels to existing facilities.

The methods and equipment currently used for remote computer input are only in their initial stages of development. Fortunately, the primitive state of remote input technology has not, to date, seriously hindered the application of computers to business data processing. The majority of commercial computer applications presently involve conventional accounting and control systems not requiring the use of high speed data accumulation or dissemination. It appears, however, that many proposed business data processing applications involving the nebulous concepts of projection and real time control are awaiting the development of more advanced remote input equipment and techniques.

Despite the problems involved in adapting systems and equipment to existing communication media, significant advances are being made in the use of remote input for commercial data processing installations. One of the most ambitious projects in this field was initiated on October 4, 1955, when the Sylvania Electric Products Company began construction of a new data processing center at Camillus, New York. The 50,000 square foot installation, expected to employ approximately 300 persons, will receive all of its data from remote locations. Hardware involved will include a large scale computer, high speed teletype equipment and 12,000 miles of leased wire communication circuits. The data flow system will be a series-parallel arrangement. Six area offices will be linked directly to the center and each area office, in turn, will be connected with from four to six sub-facilities. Prime output for this installation will be frequent summarized reports of production and marketing data with conventional financial and accounting reports relegated to a by-product status. This type of radical approach has frequently been proposed as the most profitable long-range approach to business data computer applications. Future results at Sylvania should provide an excellent opportunity to examine this promising, if unorthodox, hypothesis.

Applications

Investigation indicates that most conventional remote input applications fall into one of three categories: 1) those involving input data with a high time value, 2) applications where central processing is more economical than remote processing despite the added burden of transmission costs and 3) remote input applications involving only a portion of the total input data but allowing uniform cutoff dates despite geographical separation of reporting functions.

Typical of the first category of application where input information has a high time value are those involving source documents that affect company income. For example, the processing of summarized sales information. In one department of the Du Pont Company, regional sales offices prepare sales orders on a teletype machine connected directly to a nearby warehouse. The duplicate copy of the sales transaction reproduced on the teletype machine in the warehouse is used to initiate shipment of the order. In the sales office, a paper tape is produced automatically by the teletype machine as a by-product of this operation. It contains specific portions of the information recorded on the sales order form. In content, 15 to 20 per cent of the information transmitted to the warehouse is recorded on the tape. These sales summary tapes are accumulated and transmitted from each sales office, daily, to a central data processing installation. There, statistical computations can be made on the data to determine significant sales trends in different products, generate or modify production schedules and allocate finished goods shipment to and between the remote warehouses.

Another type of remote input application justifiable because of the speed and capacity of the central computer is that involving the processing of repetitious engineering calculations. An example of this is the use of the General Electric computer installation at Evendale, Ohio, by G. E.'s Medium Steam Turbine Department in Lynn, Massachusetts—1,000 miles distant. In December 1953, the Lynn group began transmitting turbine design data over a teletype line to Evendale for processing on an IBM 701. The computer output was later converted to punched paper tape and transmitted back to Lynn. Initially, the design parameters to be investigated were punched into cards direct from a standard data sheet furnished by the design engineer. These cards were run on a tabulating machine, certain control totals taken and a summary card punched. All the cards were then converted on a punched card-to-paper-tape device and transmitted by teletype. At the computer, the paper tape received was converted into cards and these listed on another tabulating machine to check the control totals against the summary card. The feedback of results followed the same procedure in reverse. Recently, IBM Transceivers (punch card transmitters and receivers) were installed to eliminate conversion between paper tape and cards. The resulting increase in accuracy was sufficient to eliminate the listing and check total operations. The average computer running time for each solution is about 5 minutes. Current volume is approximately fifty computer solutions per week.

The second major category of application mentioned previously were those involving remote input of data which, despite transmission costs, can be processed cheaper at a central computer. An example is the processing of construction site payrolls. A typical construction site starts off with a few men, builds up to many hundreds or thousands, and then declines to zero in a period of 12 to 18 months. The problems of mechanizing payroll procedures on location for these applications are evident. The system is either over-designed or woefully inadequate, depending upon the stage of completion of the project. The transmission of the payroll data to a central computer can produce economies for this type of application because the processing is always within the capacity of the hardware, the minimum to maximum processing time range is compressed, and the necessary accounting and control data can be easily extracted in a form suitable for subsequent home office processing. Additional economies may be realized, in some instances, by combining a number of similar payrolls into one processing run.

Equipment

The techniques now used for remote computer input can best be understood by describing remote input data origination and then following the operations back to the central computer.

Currently, commercial computer installations with high speed remote input applications use data origination devices which produce either punched cards or 5 channel punched paper tape. Despite the disadvantages of the 5 level code stemming from the necessary inclusion of the upper and lower level shifts required to handle alpha-numeric data, it has gained wide acceptance from office equipment manufacturers and users because of its compatibility with existing communications systems. The trend toward early recording of source documents in mechanical form has been promoted widely under the name of "Integrated Data Processing" or the "common language" concept.

Frequently, the data produced on common language paper tape machines has its format and content determined by the requirements of the originating operation. The sales order analysis application mentioned earlier is a case in point. In that example, an order, invoice and acknowledgement are produced on the originating device. While the volume of data extracted can be varied, the necessity for a clean format on the customers' copy precludes the insertion of zero field fills and special function symbols. Thus, the problem of unscrambling the data fields on the by-product tape is passed on to the terminal conversion equipment or the central computer.

The actual transmission of the input data, whether recorded on cards or tape, may be handled in a number of ways. If recorded on cards, duplicate cards may be produced at the terminal computer by transmission between two IBM Transceivers, utilizing either a teletype line or a voice frequency telephone line. When a teletype line is used with a Transceiver (Fig. 1), the speed averages about the same as a conventional teletype machine; six or seven characters per second. Over a telephone voice circuit, however, the maximum transmission speed jumps to 15 characters per second with up to 4 pairs of Transceivers theoretically possible on one voice circuit. When a telephone voice channel is used, the Transceiver contains a pluggable channel selector that determines which of the four allowable frequencies the machine may transmit or receive on. A program card, wrapped around a rotating drum, is used on both the transmitter and receiver for format control and to insure that all columns transmitted are received. The device has a number of self-checking features. A cursory survey of 15 of the estimated 77 commercial Transceivers in use as of November 1, 1955, indicated an extremely low error frequency. If the input data is on punched cards at the remote location but the volume of transmission too small to justify a Transceiver hook-up, the data may be converted to five channel paper tape for teletype transmission.

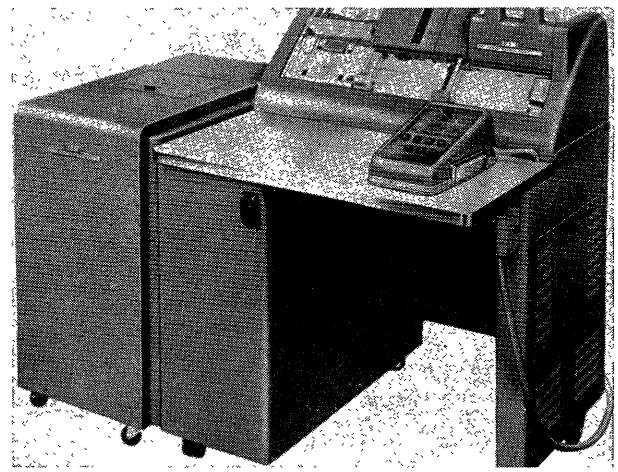


Fig. 1.

There are a number of systems for the transmission of five channel paper tape over teletype lines. Tape readers, transmitters, nonprinting paper tape punches and printing tape punches may be combined with either a common carrier wire service or a private wire. Private wire agreements involve the lease, on a monthly basis, of the lines and the transmitting and receiving equipment. Common carrier agreements may also involve lease of the equipment but alternate methods of payment for line usage are available. For example, some tariffs are on a per connection basis with a minimum charge for three minutes and an incremental charge for each minute thereafter. One common carrier charges by the word at a rate which is a function of monthly volume; another charges by the word with a day to night tariff differential of about three to one. Private wire rates are calculated at a fixed amount per air mile and vary for different scheduled periods of usage per day.

The choice of private wire or common carrier service for teletype transmission depends on volume, availability of service, maintenance arrangements and ratio of equipment cost to line cost. Frequently, a combination of private wire and common carrier service results.

Volumes of less than 300,000 characters per month are frequently more economical on common carrier services such as TWX, Telemeter, or Western Union. Above that volume, private wire service usually becomes more attractive. If the distance to the remote location is short, the actual line rental is a small part of the total system cost. For example, in a ten-mile circuit, approximately 15 per cent of the total equipment and line rental costs are applicable to line rental. The choice of service in this case is usually based on factors such as equipment offered, maintenance policy, and convenience. When the circuit extends 750 miles, however, the line costs jump up to approximately 75 per cent of total system cost and line charges become a major factor in service choice.

Conversion at Terminal Computer

When cards or tapes are received at the terminal computer, they are marked for visual identification. Transceivers are available which print as well as punch while receiving. With paper tape, a typing reperforator is normally used as the receiving device. This unit punches the tape and also types on the tape the data received. If tapes are received from a number of locations for processing during one master run, they are assembled in batches until cutoff time and, if the computer does not utilize paper tape input, converted to a form suitable for computer input. With paper tape, the most direct way to do this is with a paper tape to magnetic tape converter or a paper tape to punched card converter.

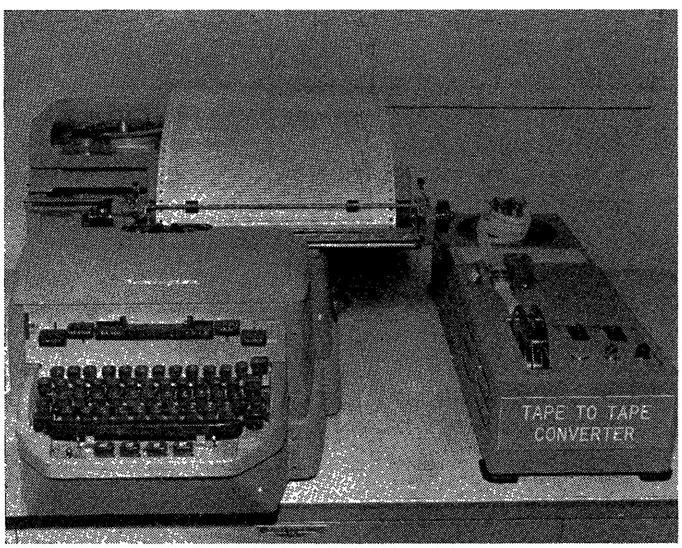


Fig. 2.

A prototype paper tape to magnetic tape converter is in use at the Du Pont Company's Univac installation. (Fig. 2). Functionally it is a keyboard to magnetic tape device, a Unityper II, connected with a paper tape reader. It produces, in addition to a magnetic tape copy, a printed copy for use as a reference document. Data is recorded on magnetic tape in blockettes of 120 characters. Provision was made for the elimination of any blockette by the inclusion of a special code symbol at the end of the data to be erased. Thus, station directing codes for the teletype network and other identifying symbols are recorded on the printed copy and then erased automatically from the magnetic tape. A flexible decoding matrix on the unit has permitted a wide latitude in determining which of the 46 nonfunctional paper tape characters shall be used with which of the 63 Univac characters.

The Sperry-Rand Corporation now has under test a high speed paper tape to Univac magnetic tape device utilizing a photoelectric tape reader. This device will permit conversion speeds, from paper to magnetic tape, of up to 200 characters per second.

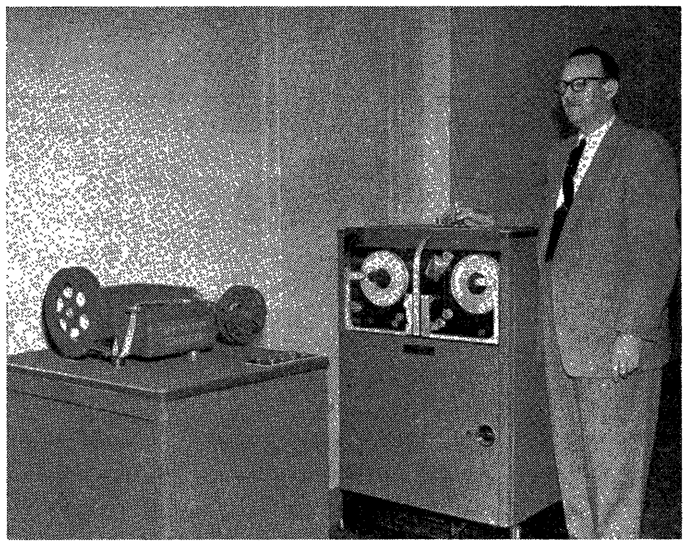


Fig. 3.

Computer output for transmission from the central computer to remote locations may be handled with a magnetic tape to paper tape converter, a card to paper tape converter or a direct card transmission on a Transceiver. An interesting magnetic tape to paper tape converter is in use by Du Pont. (Fig. 3) The device consists of a low speed magnetic tape reader connected to a paper tape punch. Format control and the inclusion of the necessary station directing teletype codes are recorded on tape during the output editing run on the computer. Figure and letter shifts are automatically punched in the output tape as required. Standard editing routines for low speed Univac output printers are used to arrange the format compatible with the receiving teletype equipment at remote locations. A commercial version of the Univac magnetic tape to paper tape converter is expected to be available in 1956.

An alternative to tape or card conversion at the terminal computer may become practical with a self-checking magnetic tape transmitter and receiver about to undergo field trials. This Sperry-Rand device reads information from magnetic tape at a remote location, transmits the data over a voice frequency telephone circuit at a rate of 150 to 200 characters per second and reproduces the data on magnetic tape at another transmitter-receiver on the terminal end of the wire.

Accuracy

The facilities of the communication utilities are not designed to provide the low undetected error frequency found in modern computing systems. Lacking self-checking transmitting and receiving equipment, it has become standard practice to design the data flow system to cope with the expected errors. The degree of additional complexity this builds into any given application depends upon the possible consequences of undetected errors in that application. The resultant system is usually built up of document counts, sequential numbering systems, control totals and program checks in the processing instructions. In this latter case, the expected occurrence of numeric data, instead of alphabetic, in a numeric field is tested, non-existent product code numbers are brought to the operator's attention, payroll checks above a maximum expected amount are noted and other logical choices executed.

In 5 level paper tape systems, experience has indicated that the random mutation or dropping of a single character is rare. Most system errors destroy a string of characters and thus increase the probability of the bad transmission being detected. The single numeric substitution error in a numeric field not under a control total, while rare, generally causes the most concern. These are seldom as probable as might appear, however, considering that only 9 of the possible 32 mutations would result in erroneous numeric input.

To qualify as an undetected error, a teletype mutation must penetrate visual inspection, field distortion checks on conversion equipment, logical computer program checks and batch, sequence and control total verification. Errors which do not slip through the system frequently present more of a problem than those that do. If detected before computer processing, retransmission of the data batch is adequate. Errors detected during processing cause all data associated with the transmission to become suspect in view of the error grouping phenomenon mentioned earlier. In some cases it is feasible to have error correction routines available at the computer which will back-out of the master file all previously processed input associated with the suspect transmission.

Self-checking data transmission and receiving equipment eliminates many of the remote input problems of the system designer and programmer. The IBM Transceiver is the only such device currently available. Each character read from the card to be transmitted is encoded into a four-out-of-eight pulse train. At the receiver, the impulses are checked to see that they are a legitimate four-out-of-eight representation. If an invalid character combination is received, an error condition is set up in the receiver. The next request-for-card-release signal, sent by the transmitter at the end of each card, is denied by the receiver and operation stops. After recognition of the bad transmission by the operators on each machine, the card is re-transmitted.

One major difficulty involved in designing teletype equipment with built in checking is that the five level code itself offers no easy way for the inclusion of parity checks to detect dropped or mutilated characters. Currently there are a number of organizations working on systems to permit the inclusion of checking procedures without unduly complicating the encoding or decoding functions. For example, one large teletype user is developing a simplified system for teletype equipment by which a parity check is made along each of the five channels on tape. When the carriage return and line feed functions are performed on the originating teletype machine, the parity balance for each channel is punched in the tape. At the receiving machine, the carriage return and line feed symbols alert the system to decode the next character as the parity check. It is understood that the Bell System is working on a similar technique which automatically scans the bits in each channel and includes the parity check punches after a fixed number of characters. Western Union is reported to be working on a system which transmits each bit twice, polarized first in one direction and then in the other.

Remote Programming

The definition of remote programming used here refers to the preparation of computer instructions by personnel geographically separated from the computer installation. It implies that a remote location would transmit detailed coding, pseudocode or the problem description to a central computer where a processing run would be scheduled and the results transmitted back to the source. Input data for the run could come from either the remote location or master files maintained at the computer.

Currently, there is little demand for this technique. The noise about central data processing for better management control has not yet subsided enough for decentralized operating units to be heard. Current organizational thinking indicates that decision making should be executed at the lowest possible level in an organization. It is hoped that the next major area of business data processing application will be directed at supplying better information to decentralized management functions. One objective might be to enable each remote function to better relate its operations to those of the organization as a whole. Central data processing with remote input-output facilities and adequate remote programming technology could make this feasible by enabling remote interrogation of master files at the central installation. This type of application represents a logical projection of the current trend in data processing.

The problems of remote programming differ from those of conventional programming only in order of importance. The problems of remote debugging shift the emphasis to the reduction of error probability and development of adequate diagnostic routines to enable easy detection of programming errors which do occur. Experiments with remote programming, utilizing extensive program diagnostic routines, have been successfully carried on by General Electric's Medium Steam Turbine Department. Results indicated that the diagnostic routines and improved desk checking more than compensated for the absence of programmers from the computer by decreasing expected debugging time.

Automatic coding techniques, which reduce the number of instructions written to produce a given program, should facilitate remote programming by reducing error frequency and minimizing programming effort. The difficulties involved in debugging current automatic routines when errors do occur, however, indicates the desirability of including adequate diagnostic routines to provide the remote programmer or operator with pertinent error detection and correction data. The promised development of one automatic coding routine utilizing verbs and nouns as pseudocode input could be a significant factor in facilitating remote programming. This approach would provide a communication channel enabling the remote programmer to specify precise program requirements in a verbal flow chart meaningful to those who process it.

Conclusion

Remote input applications for business data processing installations are feasible with existing technology. The shortcomings of current conversion or transmission equipment can be circumvented with proper data system design.

Data utilized in remote input applications is frequently a by-product of an operation which governs its format and content. As source document mechanization spreads, through the paper tape "common language" concept, more data in a form compatible with transmission facilities will be available for the improvement of central accounting and control operations provided unnecessary function codes are not required by the terminal conversion equipment.

Early remote input applications indicate management is becoming aware that present business reporting techniques which indicate change over a fixed time period must be supplemented by techniques which indicate the rate of change. As this awareness penetrates, it is reasonable to expect a rapid growth in this area.

Remote data transmission is a two-way street. The benefits derived from the rapid accumulation and processing of source documents for centralized reporting should be compounded by utilizing the same facilities to make the resultant data available for better decentralized operation.

Discussion

L. Difford (Department of Defense): Would you care to discuss the use of a central computer and remote inputs for obtaining quickly a national radio or television index?

Mr. Fitzgerald: I would weigh very carefully the feasibility of high speed transmission and processing in a case like this. Perhaps the major problem here is the mechanical recording of large amounts of source data over a short period of time. It might be possible to work out a system by which the viewers could be called on the telephone, receive a recorded message and requested to dial "yes," "no" or a TV channel number. The resultant impulses might then punch a tape for subsequent automatic processing, much like the communication utilities do with their automatic message accounting systems.

Lt. Col. McPerson (U. S. Army): What, in general, is the state of the art of "on line" transmission of data?

Mr. Fitzgerald: I have had little contact with "on line" processing of business data. It appears that such applications are currently limited to special purpose banking, inventory and reservation equipment.

R. E. Wright (Teleregister Corp.): Would you discuss error detecting and correcting features of present equipment designed to reduce system faults due to transmission; that is, line hits and repeater errors?

Mr. Fitzgerald: I presume you are referring to teletype networks and voice circuit transmission using Transceivers. I know of no teletype equipment with either error detecting or correcting features. Errors are usually detected by checks built into the data flow system and corrections made by re-transmission. The IBM Transceiver detects mutilated characters by use of a self-checking four-out-of-eight pulse code. Matched program cards on the transmitter and receiver detect dropped characters. Correction is made on the spot by the operator by retransmission of the card in which the error occurred.

P. W. Larsen (Teleregister Corp.): Would you please comment on developments for direct input from remote points without going through tape or punched card mediums; for instance, keysets used for both input and output which are directly connected to a computer via leased circuits?

Mr. Fitzgerald: This "on line" area of business applications currently appears to be limited to special purpose computers utilizing muliplexing on input and output. Most remote input applications of this type use teletype lines and paper tape buffers. I know of no keyboard input devices on special purpose magnetic drum machines located more than a few hundred feet from the central hardware. Perhaps few commercial applications have arisen which can economically justify utilizing long line leased circuits for the relatively slow input of direct keyboard entry.

Lt. H. S. Brun (Signal Corps Engineering Laboratory): Why not interconnect computers through buffers rather than slow speed terminal and conversion equipment?

Mr. Fitzgerald: Actually the paper tape which necessitates the use of conversion equipment is a buffer. It takes up the slack between data origination speed and efficient computer input speed. Unfortunately, the data stored on paper tape usually does not have the same pulse code representation as that of the central computer. This leads to the use of conversion equipment which increases the access time to the data. Computers could be linked together over leased lines with more efficient buffering, such as magnetic cores, if the input-output speeds were carefully balanced and their programs synchronized.

H. Ilger (Rand Corp.): What steps are being taken to eliminate the volume of expensive converting and/or intermediate equipment?

Mr. Fitzgerald: Unfortunately, there appears to be a trend toward building more logical operations into paper tape conversion equipment. In some cases, it is necessary for the data originator to program function codes for converter operation into the input data format. This approach should be balanced against the possibility of passing format control on to the central computer by performing only code conversion on the converter.

G. Licht (Wayne University): Do you think that the present shortage of transmission facilities (lines) will seriously affect the creation of far-flung computer systems?

Mr. Fitzgerald: I understand only real shortage of transmission lines is in the area of high quality voice circuits and broadband circuits for facsimile transmission. The waiting periods for procuring leased circuits have been shortened considerably during the past year. The cost of leased lines for remote input rather than circuit availability appears to be the determining factor in most applications.

G. E. Reynolds (Air Force Cambridge Research Center): Please expand on the use of verbs and nouns in diagnostic routines.

J. A. Sperling (Western Union): Pseudocoding using verbs and nouns as checks of accuracy; can you give us a simple example of how this is done?

Mr. Fitzgerald: Automatic coding techniques permit the programmer to write computer instructions in a simplified form called pseudocode. The computer is used to convert these into legitimate instructions. Conventional pseudocodes are meaningful to the remote programmer but do not define the problem to a person at the central computer. The use of verbs and nouns as pseudocode will permit the remote programmer to specify exactly the operations desired in a discrete form. This should decrease error probability by reducing the volume of instructions transmitted and allowing logical verification by the central computer staff.

J. Smith (Smith Mfg. Co.): How many characters per minute can be transmitted on one Transceiver? If you used enough teletype machines to occupy a full voice channel, wouldn't the output be more than the four Transceivers?

Mr. Fitzgerald: Approximately 15 characters per second can be transmitted over one Transceiver, with up to four pairs possible on one voice circuit for a total of 60 characters per second. I understand that the communication utilities can put up to sixteen teletype channels on one broadband circuit for a possible 160 characters per second. The necessary repeaters and band splitting equipment, however, pushes the rental of sixteen teletype circuits up to about five times the rental of a voice circuit.

F. J. Cahill (Arabian American Oil Co.): Does tape to tape converter automatically code from five channel to six, seven or eight channel and vice versa?

Mr. Fitzgerald: The paper tape converter shown (Fig. 2) converts only from five level paper tape to seven level Univac code. Likewise the other one, the magnetic tape to paper tape device (Fig. 3) converts only from seven level Univac code to five level teletype code.

D. R. Klusman (American Telephone & Telegraph Co.): At what monthly volume of data do you think serious study of mechanized data transmission facilities is warranted?

Mr. Fitzgerald: I would say 200,000 characters per month.

P. J. Ullman (New York City Housing Authority): With 100 window posting locations in various parts of New York City, covering 75,000 accounts, what system would be indicated for centralized accounting?

Mr. Fitzgerald: Window posting machines are available which create a byproduct punched paper tape. These tapes could be mailed, or transmitted by teletype, to a cental accounting office. Paper tape to card converters or a paper tape to magnetic tape converter could be used to convert the data into a form suitable for central data processing equipment input.

J. S. Woldringh (Shell Research Laboratory, Amsterdam, Netherlands): What kind kind of checking possibilities are available to insure that no errors are introduced in the wire transmission stage, such as in the case of radio transmission?

Mr. Fitzgerald: I suspect that the principle of Transceiver checking is applicable since there is now a Transceiver network operating between Morocco and Washington, a 3,000 mile span, over a radio channel. If they are using conventional Transceivers, each pulse burst is checked for a legitimate four-out-of-eight code and correlated with a column on the master program cards. After the last card column, the transmitter sends an end of card request pulse to the receiving transmitter. If everything is acceptable and the two master program cards are in register, the receiver automatically sends back a card release signal.

Developments in Programming Research

C. W. ADAMS†

THE MEN AND WOMEN who design the kitchens of tomorrow, the new automobiles, or the new office machines aim at functional as well as esthetic appeal. They strive to improve functionally in two ways—by *mechanization* and by *systematization*. These are the two basic approaches to the simplification of the tasks of life. These, then, are also the two paths to be explored and developed by programming research.

Unfortunately, the sensuous delight which I personally derive from enunciating such polysyllabic words as mechanization, systematization, and (best of all) automatization, seems not to be shared by others. I suppose I must, therefore, beg Noah Webster's forgiveness and keep myself up with the times by dropping a syllable or two. Consider with me, then, what are the possibilities for the *automation* of computer programming through *mechanation* and through *systemation*. Investigation of these possibilities comprises the field of programming research—or P/R as it might be called.

† Westinghouse Electric Corp., Philadelphia, Pa., Assistant Professor of Digital Computation on leave from M.I.T., Cambridge, Mass.

Mechanation

In order to mechanize anything, one must first systematize it. One of the greatest stumbling blocks in programming research is the obsessive urge to mechanize. This seems to drive many researchers to attempt to mechanize procedures which have not yet begun to be systematized.

Of course, there are certain aspects of computer programming which lend themselves well to mechanization, and much effort has been spent in developing programmed techniques to deal with these.

In the early days, program research was largely concerned with what now seem like very rudimentary problems indeed. Strange as it seems, many computers were nearly completed before anyone even established how they would load instructions into the storage element to get the problem started.

Mechanation consists of transferring from men to machines as much of the effort, both physical and mental, of a given task as can economically be handled on the machines. Thus, simplifying the coding of a computer program by making the computer perform as much as possible of the clerical effort is mechanation of coding. That part of the mechanation which is achieved by computer designers is called "logical design"; the part achieved by programmers is called "automatic coding."

Automatic Coding

The ACM Glossary defines automatic coding as "any technique in which a computer is used to help bridge the gap between some 'easiest' form of describing the steps to be followed in solving a given problem and some 'most efficient' final coding of the same procedure for a given computer."[1] I have heard two major criticisms of this definition—one that it is too broad, the other that it evaluates the techniques a priori. Stephen E. Wright, in a sparkling d-"Evaluation of Automatic Programming"[2] presented at the ACM Meeting in September, 1955, said of the definition: "in effect it says that automatic coding is any *good* technique of problem preparation using a computer." He suggests, "as a working definition," that "automatic coding translates instructions that a computer cannot execute into instructions that a computer can execute." This is usually done by "a routine executed by the computer itself, although a smaller computer, a coding machine, or punch card equipment may be used."

Whatever the definition, the bulk of the work on automatic coding up to now has resulted in routines which perform some variation of one or more of these six functions:

1. *Conversion to Binary* of decimal data, decimal instruction addresses, and mnemonically-written operation codes. This is a minimum function for binary computers to have and is one which built-in decimal arithmetic virtually obviates, although even decimal data must usually be edited from convenient input format into the form desired for machine processing.

2. *Subroutine Adaptation* by means of relative addresses and preset parameters. (This function as well as conversion to binary is performed in the EDSAC at Cambridge by only forty-one "initial orders"[3]—the loading routine which was written more than six years ago by David Wheeler and which still stands as the leading example of programming virtuosity.)

[1] "First Glossary of Programming Terminology," Association for Computing Machinery, 2 E. 63 St., N. Y. 21, N. Y., p. 5; 1954.
[2] Paper not yet published.
[3] M. V. Wilkes, D. J. Wheeler, and S. Gill, "The Preparation of Programs for an Electronic Digital Computer." Addison Wesley Press, Cambridge, Mass. pp. 15–21, 159–162; 1951.

3. *Storage Allocation* of programs written with symbolic addresses and literal addresses. In these a programmer turns over to the computer the task of assigning instructions and data to specific storage locations. Instead, he indicates the operands in his instructions by freely chosen symbolic addresses or, in the case of small constants, by writing the number itself in place of choosing any actual or symbolic address. Symbolic coding, which was first described independently by Maurice Wilkes of Cambridge University and by Nathaniel Rochester of IBM, is a feature of almost every new automatic coding system. Literal addressing, which had been suggested by David Wheeler while he was at Illinois, was rediscovered by the developers of the IBM 702 Autocoder and is included in that system.

4. *Program Latency Reduction.* In a serial storage system such as a magnetic drum, the speed of processing can often be greatly increased by judicious selection of the locations in which individual instructions and data are stored. This is often called optimum programming, a term which seems not to describe the process or its results but merely the wish behind it. Because latency reduction is a cut-and-try process, Barry Gordon of Equitable Life recently made the interesting suggestion that a large computer, with its greater computing efficiency and auxiliary storage facilities, could be used to do a better job of preparing a reduced latency program for a drum computer than could economically be done on the small drum computer itself.

A different sort of latency problem arises when a program is too long to fit in primary storage, be it cores or drums, but must be stored on a slower, serial auxiliary, either drum or tape, and then brought into high speed storage one segment at a time. Mechanized procedures for dealing with this storage hierarchy problem are being developed at Remington Rand, M.I.T., and elsewhere. However, the greatest promise seems to rest in the FORTRAN system being designed under John Backus at IBM.

5. *Translation of Macrocoded Instructions.* This is the area most widely identified with automatic coding. The translation starts with the decoding of instructions written in some macrocode. The term pseudo code is more commonly used, but I tend to favor macrocode as being more meaningful. So far as I know, this term, like the term literal address, originated with the developers of the IBM 702 Autocoder. In any event, the decoding operation determines which library subroutine, with what values of parameters, is needed to make the computer perform the equivalent of the macrocoded instruction. Then the required subroutine is selected and adapted to the situation. In some cases, the subroutine is generated, a technique which is very useful but not as exciting as it seems. Generation is merely a special case of subroutine adaptation in which the subroutine is un-

packed, unwound, or otherwise expanded from some skeletal form to its final form.

Let us return to the matter of selection. Once a subroutine has been selected, it can either be performed on the spot, before the next macrocoded instruction is translated, or it can be recorded to produce a machine-coded program to be performed later.

In the first case one has an interpretive routine, which produces numerical results; in the second case, a compiling routine which produces programs. The difference between interpretive and compiling techniques is entirely one of economics. When macrocoded instructions are repeated cyclically within a program or when a program is reused, the interpretive routine must retranslate whereas the compiling routine need not. On the other hand, interpreted programs require less storage than compiled ones do. The original concept of an interpreter, in fact, was as a means of saving storage space at the expense of computing time by condensing coded references to commonly needed subroutines into as little space as possible. Now that larger internal storage elements, more auxiliary storage, and more elaborate built-in operations are available, the interpretive routine seems to be on its way out. For example, the two most important features of Speedcode and the other interpretive routines written for the IBM 701, namely floating point arithmetic and address-modification facilities, have been built *into* the IBM 704. In demonstrating the usefulness of such facilities, the interpretive routines have served us well. Their brief careers have not been in vain.

The success of a translator, be it a compiler or an interpreter, depends largely on the macrocode which it is designed to tranlate. In choosing this code or language, the program researcher is faced with much the same problems that beset the designer of a computer. What facilities will be useful? Will each be worth the cost? Before these questions can be answered one must know who is to use the language. Will it be for engineering computation or for business data processing? Will its primary purpose be to reduce the effort required of an experienced coder who is on intimate terms with his computer? Or will it be intended to simplify the task of the novice or the occasional coder? Generalized Programming, as developed under Dr. Mauchly at Remington Rand aims at the experienced user. The Summer Session Computer[4] at M.I.T. and the EASIAC at Michigan are for training purposes only. Algebraic Coding,[4] originated by Laning at M.I.T., appeals to the occasional coder, while FORTRAN at IBM and A-2 at Remington Rand are intended to satisfy the occasional user and the experienced one as well. B-Zero, which like A-2 is the product of Dr. Grace Murray Hopper's group in Remington Rand, extends some of these ideas into a form more useful in business problems, while Dr. Mitchell's BIOR is aimed at the most important problems of business coding that seemed capable of mechanization (input, output, and rerun routines), and is intended for the professional business user.

One other purpose of translation has been less thoroughly explored, although a number of ideas have been described by Dr. Saul Gorn[5] at Aberdeen Proving Ground. Let me read part of a letter I received from E. F. Cooley, Director of Methods Research for The Prudential Insurance Company, which has over one hundred people engaged in planning and programming for IBM 650's and for seven IBM 702's and 5's. Says Mr. Cooley, "we are interested in any method anyone may devise for conversion of the machine instructions for one computer to the proper instructions for a different machine. When you included this sort of thing in your M.I.T. classes, we thought of it as an interesting trick but perhaps of little importance. But it is a sure enough problem under our present circumstances where we will program the same work for both the 705 and the 702 and a good portion of it for the 650 also." Mr. Cooley's problem is not likely to be solved soon. Such translations are simple in principle but extraordinarily difficult in practice when one attempts to preserve or increase efficiency in the face of various word lengths, instruction codes, tape facilities, and other logical differences in the machines.

6. *Aids to Mistake Diagnosis.* Locating mistakes that planners and coders inevitably let creep into their programs is something that everyone talks about but that very few do anything about. But, they say, compiling or interpreting reduces the likelihood of mistakes by reducing the number of instructions to be written and by eliminating the detailed coding of input, output, loading and other relatively straight-forward routines in which most of the careless mistakes are made. This is true, but it is not enough. One of Mr. Wright's criticisms of automatic coding, while ill-founded in principle, is all too valid in practice: "There is little doubt that actual addresses are far more convenient than symbolic addresses in error detection." Program researchers should hang their heads in shame over the number of programmers who make full use of powerful symbolic address and macrocode translation facilities in preparing their programs, then join Mr. Wright behind the buttons and indicator lights in extravagant, horse and buggy techniques for debugging these same programs. Certainly it is difficult to apply manual debugging techniques to programs prepared using automatic techniques. To be effective, automatic coding must go all the way, not

[4] "Symposium on Automatic Programming for Digital Computers," sponsored by NMCAP, office Of Naval Research; May 13–14, 1954; proceedings available through the Office of Technical Services; pages 40–68 (all 16 papers discuss various coding systems for handling scientific and engineering computations).

[5] *Ibid.*, pp. 74–83.

stop short where the going gets rough.

The proper way to debug a program, once it has been carefully written and at least casually checked by hand, is to put it on the computer, note all the symptoms, and take the program back to the desk for diagnosis and correction. Such a procedure requires that there be an efficient way to put the problem on the computer, record any data which might help in the diagnosis, and get off. With regard to the recording of symptomatic data, selective tracing and changed word post mortem routines have been in very successful use here and there for years. Of course, it goes almost without saying that these routines should perform a retranslation from machine language to the symbolic address and macro-coding notation employed in writing the program. Tracing and post mortem retranslation techniques are rather difficult to develop when auxiliary storage is employed, especially when programs are segmented on a drum or tape, but these difficulties should not be insurmountable.

The quick change from one program to another is not possible when numerous switches must be set, buttons pushed, and tapes mounted, dismounted, and labelled. But switches need not be set, numerous buttons pushed, nor more than one tape mounted or dismounted if the proper program was available to get things started during a trial run. Switches, in fact, should never be used. Push-buttons may often be the best means of communicating to the program just what variation is desired, and are indispensable for starting the computer in the first place, but program alteration should be programmed and be controlled from punched cards, tapes, type-ins, or push-buttons, not from switches which may be left in the wrong position when the next program is run. All of the test data which is to be generated or has been prepared in advance can be on one tape or one deck of cards for the trial runs. A supervisory program can then, working from specifications from the single input, disperse data out to the proper tapes and otherwise prepare for the actual trial run. Logging and recording of symptomatic data should also be automatic.

Not only during trial runs but also during actual operation should the routine control of the computer be given to the program rather than to the operator. Bruse Moncrieff at RAND Corporation is working on a routine for the 702 to accomplish this. Says Bruse, "following the rule that we ought to mechanize first those aspects of the total situation which either cost the most or annoy us the most, I have turned my attention to the problem of the day-to-day operation of an automatic data processor. The things that annoy programmers the most are operators, so I am attempting to all but program him out of existence. There are certain phases of his work, mostly involving manual dexterity, which of necessity have been preserved. I have tried to remove all the thinking from his job, since this is what people do least efficiently. I like to think of this proposed routine as an automatic supervisor rather than operator since it will be telling the human operator what to do."[6]

Apology

At the start of this paper, I indicated that to me programming research is the investigation of ways of automatizing the preparation of computer programs through systemation and mechanation. Because of the preoccupation that all program researchers have had with mechanation, I have thus far devoted my remarks entirely to a brief survey of what has been done and what might yet be done in this area. In this survey I have mentioned the names of some computers, coding systems, and people. I trust that those computers, systems, and people will not feel unduly flattered and that the many I have had to omit will not feel unduly hurt.

One Package

It has taken longer than I would have imagined, but the computer manufacturers are at last realizing that they need not teach their customers the innermost workings of their computers, not even the details of the beautiful excess-three or the diabolically clever biquinary code they have so ingeniously used. By the same token, let us treat the work of the program researchers like that of the logical designers. Let us offer the customer a package, complete with techniques for using it. Let us no longer act as if basic automatic coding techniques were not a part of the system. If the customer *wants* to modify the automatic routines, by all means let him, just as we will let him modify the electronic or the mechanical features if he wishes. But give him the option of taking the system in a package, as is.

Systemation

To make the package acceptable, the designers and the program researchers must do a thorough job of systematizing the use of their computer and of mechanizing as much as can effectively be mechanized about its use. This requires much more systemation and much less mechanation than we are developing today. More powerful machines open up great new possibilities for using them; we cannot waste time in obviously misdirected or repetitive programming effort. Let us, for the business problems of today, mechanize only those housekeeping functions which can be made routine today. Once we devise a truly systematic approach, the question of what else to mechanize and how to mechanize it will become easy to answer. Let us therefore turn our most talented researchers now to the great task of *systematizing* the job of planning, coding, debugging, and operating a program.

[6] Letter soon to appear in "Computers and Automation."

Discussion

R. A. Kirsch (National Bureau of Standards): Do you think that the possible future use of large fixed memories or slow memories combined with very small, fast memories will cause an increased interest in interpretive techniques?

Mr. Adams: Not particularly so. I think there is a great potential for research into the effective use of storage hierarchies (*i.e.*, making use of large slow memory facilities in conjunction with small, fast ones). Possibly large fixed storage would encourage use of interpretive techniques, but large erasable storage would very likely lead to compiling techniques just as it led the Univac people there in the beginning.

B. Colten (Datamatic Corp.): Are you familiar with any current research toward self-diagnosis procedures between prepared programs and computing systems?

Mr. Adams: I am afraid I don't know what self-diagnosis would be other than to have the automatic routine detect any obvious errors that the coder had made. This, of course, is part of what one gets when he uses symbolic addresses and psuedo-codes, both of which provide a fair amount of redundance that can be and should be checked as a matter of routine.

J. O. Pederson (N.S.A.): What is the minimum auxiliary memory needed to apply to automatic programming techniques?

Mr. Adams: That is a good question. Automatic techniques originated on computers having storage for only about 250 words total. Therefore, the minimum would be quite small. On the other hand, most present day automatic coding systems use routines of between 4,000 to 20,000 coded words of instructions and data.

J. F. Rafferty (National Bureau of Standards): At the SEAC computer installation at the Bureau of Standards the programmers were interested in devising a way to record the engineers on magnetic tape, read them into the machine, and press the "Memory Clear" button. Conversely, as an engineer, I have been interested in reports such as yours on the attempt to program the programmers out of business.

Storage and Retrieval of Information

LOUIS N. RIDENOUR†

THE HISTORY of the development of modern information-processing machines is largely the history of the development of competent memory systems. These must be capable of storing digital information for longer or shorter times, of making desired information available with no more than a given delay, and of having total information-storage capacities suited to one or another demand.

Because memory-system design is so fundamental to the design of information-processing machines—and to systems which use those machines—I intend to subordinate my discussion of memory-system hardware to a discussion, which I think is long overdue, of the interaction between machinery and methods. It is customary to believe that an operational requirement calls into being the equipment and devices that are needed to meet it. This always happens to some degree, and may occasionally occur in a fairly pure form. Far more often, however, our thinking about ways in which a given information-processing job might be done is unduly constricted by the availability of what appear to be suitable system elements.

Now it is perfectly true that the development of the individual components making up an information-processing system is a long and expensive task, not to be undertaken lightly. There is always a major compulsion to make do with what is available—especially since system design, if it employs available components, is a pencil-and-paper exercise, cheap to conduct and easy to alter. Nonetheless, I have a feeling that the failure of system designers to demand the components they would really like to have may be leading to a tendency of the component designers to provide, not the components which system designers really want and need, but rather those which seem easy to realize in the current state of the art.

The foregoing statements are rather sweeping, and it will be best for me to attempt to support them in specific terms. I shall do this by discussing memory-system design. First of all, let me distinguish among the major categories of memory systems, as they appear today, so that the discussion can be still further specialized to only one type of memory.

The earliest memory requirement to arise, and the one that is absolutely fundamental to the realization of any sort of information-processing machine, is that of providing the operation registers of the machine. In the logical unit of a machine there are ordinarily several specialized storage registers which are used to hold computer "words"—that is, numbers, instructions, or alphanumerical entries—which are to be compared, added, shifted, or otherwise operated upon. It is possible to design a purely serial machine in which most or all of these operation registers are indistinguishable from the storage registers of the inner memory, but this is something of a *tour de force*, since any such serial machine can be very greatly speeded up by providing it with a few specialized operation registers.

† Director of Research, Missile Systems Div., Lockheed Aircraft Corp., Van Nuys, Calif.

In the ENIAC, which was the first really high-speed electronic digital computer, great attention was paid to the design of the operation registers. They were composed of ten-element ring counters, each element being a vacuum-tube toggle. Numbers were carried in the registers as arrays of decimal digits. While detailed design changes have occurred since the realization of the ENIAC, its operation registers are really not so very different from those used in more recent machines. The vacuum-tube toggle (no doubt soon to be supplanted by the transistor toggle) is still the basic element. Though this is a relatively expensive way of realizing a storage system for digital information, its use is justified by the fact that the total storage needed in the operation registers is small, while a considerable premium is placed on high operating speed.

As we proceed through this catalog of information-storage systems, the cost per bit of stored information will fall, in consequence of the fact that the required storage capacity is rising. Concurrently, the operating speed of the memory will decrease. The next class of memory system, in this hierarchy, is the working memory internal to the information-processing machine—what von Neumann called the "inner memory."

This is an array of storage registers, each capable of storing a computer word. The operating speed of the inner memory, which is measured by the average time needed to obtain from it a word selected at random, is rather less than the required operating speed of the operation registers. Nevertheless, this speed must be high enough to be in reasonably good balance with the speed of the operations carried out in the logical unit. Speaking very roughly, this means that about half the total operating time should be used in getting words out of the inner memory and putting them back again; the other half is used in carrying out the actual logical manipulations.

For the purposes of our present discussion, it is interesting to note that in the ENIAC, as it was originally designed, the auxiliary storage registers forming the inner memory were realized in exactly the same way as the operation registers were. They were also arrays of vacuum-tube ring counters. Since registers of this construction are extremely complicated and expensive, there were only 20 of them provided in the original machine.[1] It soon became clear that a larger inner memory was required, and that a sufficiently large one could never be provided by copying the design of the operation registers, because of prohibitive complexity and cost.

The next generation of machines following the ENIAC was mainly patterned either after the UNIVAC or else after the machine built by von Neumann and his colleagues at the Institute for Advanced Study. In both these types of machines, a major design innovation was the provision of an inner memory of greatly increased total capacity. This made possible the realization of the stored-program concept. In machines of the UNIVAC type, the larger inner memory involved the acoustic-delay-line type of information storage; in machines of the "Institute family," use was made of the electrostatic storage schemes pioneered by F. C. Williams and his co-workers at Manchester.

Both the new schemes for realizing inner memory met the requirements of lower cost per stored digit than the cost of the vacuum-tube toggles of the arithmetic registers; both, unfortunately, had severe disadvantages which are now leading to their replacement. The delay-line memory is either quite expensive or else rather slow, though it can be made to be highly reliable. The electrostatic memory is far cheaper for a given size and operating speed, but difficult to make adequately reliable. In either case, the limitations of the inner-memory hardware have had a tendency to restrict the total size of existing inner memories to capacities which are rather too low to permit the realization of proper balance in applying machines to many common problems.

Fortunately, there is now coming into universal use a type of inner memory system which appears to have quite satisfactory properties. This is the coincident-current magnetic-core memory system which was pioneered at M.I.T. and at RCA. While the operating speeds required of an inner memory demand that the magnetic-core memory be provided with rather expensive vacuum-tube drivers, the magnetic-core memory is cheaper than earlier types of inner memory in any realistic size, is very much cheaper in larger sizes, and is tremendously reliable.

While there are other developments aimed at the realization of inner memory systems, notably the exploitation of ferroelectric storage schemes, the magnetic-core memory promises to be adequate for quite a time to come.

Even when a modern information-processing machine has been provided with an inner memory of adequate size, it is still imperfectly matched to its input and output organs. Its internal speed of operation is usually too fast to enable the machine to be matched directly with readers of punched cards or punched tapes, common input devices; or with card or tape punches, or printers, at the output. Because of this, and because the total size of inner memory systems has been less than ideal on account of the limitations of earlier types of inner memory, it is now conventional to interpose some type of magnetic tape or drum storage between the world of men and the world of the computer.

The magnetic drum involves an average random-access time which is equal to the time taken for half a revolution of the storage surface. Generally speaking, this time is about a thousand times longer than the

[1] It is true, but beside the point, that additional registers which could be consulted by the machine were provided in the form of "function tables." These were arrays of switches which could be set once for all at the beginning of a problem. They were not "inner memory" in the proper sense, since the machine itself had no capability to alter the entry in any function-table register, only to consult it.

random-access time typical of a modern inner memory. However, the rate at which information can be read from tapes or drums, once the desired information has been located, is comparable to that possible in reading into and out of the inner memory. Thus a reasonable balance between buffer memory and inner memory is usually achieved by reading information in sizeable blocks from a magnetic drum or tape to the inner memory, or *vice versa*.

In the foregoing historical sketch of memory-system development, we have touched on three rather different requirements: the operation registers of a machine, its inner memory, and the input-output buffer. In all three cases, the current state of technical development can offer adequate means of meeting the requirement. Operation registers are realized in terms of vacuum-tube (soon transistor) circuits. Magnetic-core memory systems appear to be ideal for inner memory. Magnetic drums and magnetic tapes serve the buffer role quite well, though there appears to be a possibility that magnetic-core arrays using relatively cheap drivers, as can be done when the random-access time is lengthened, may be competitive in this application.

In addition to the types of memory system which have thus far been discussed, one more kind is important. It is especially important to the application of information-processing machinery to the requirements of business and industry. This is the very large file of current information, whose use and management may be the chief task of an information-processing system. In contrast with the other types of memory, this type has not yet been properly mechanized.

Let us first define what is meant by "large." Operation registers ordinarily total a few hundreds, or at most a few thousands, of bits. Inner-memory systems may involve the storage of a few hundreds of thousands of bits; input-output buffers, as we have defined them above, generally store a few millions of bits. The large file, on the other hand, must hold something in the order of 10^9 to 10^{10} bits. The subscription file of a mass-circulation magazine, for example, may contain 1,000 bits of information for each of five million customers.

This total quantity of information is sufficiently great to make major requirements on the storage means. To date, attempts to meet these requirements with actual hardware have not been too successful. Just as the design of the operation registers was carried over into the realization of the inner memory in the original ENIAC, here it has been the case that the design of input-output buffers has been pressed into service to meet the needs of the large file. Thus devices using magnetic tape have been widely proposed.

The information density which is possible with magnetic tape is not very high. Current practice is to record perhaps 200 bits per inch along a track, and to space tracks about ten to the inch. Thus we have some 2,000 bits per square inch, so that a file of 10^{10} bits will require 5 million square inches of magnetic recording surface, or 130 miles of half-inch tape. This enormous quantity of tape creates access-time problems; attempts to meet these problems are responsible for the invention of the Tapedrum built by Clevite-Brush and of the MASS by the Telecomputing Corp., which latter device uses thousands of individual lengths of magnetic tape, each with its own reading and writing means.

While such devices reduce somewhat the disabilities of magnetic recording to cope with really large files of information, I think it is quite apparent that a wholly different storage scheme is required. The use of photographic means of storing digital information has been proposed, and some development work has been carried out, notably at Eastman Kodak[2] and at the International Telemeter Corp.[3] When suitably mounted to permit adequate registration in the reading process, a photographic emulsion can readily store more than a million bits per square inch, so that a file of 10^{10} bits uses only 10^4 square inches of recording surface, or an area equivalent to that of a few hundred punched cards.

One of the objections that used to be urged against the use of photographic recording of digital information is that this type of recording is not erasable, in the sense that magnetic recording is. It has become clear that this objection is not well founded. For many types of records, permanence and insurance against accidental erasure or alteration are very important virtues; so is the ability, which photographic recording provides, to enable ready human reading of any part of the record. Changes and additions to the file, which for any given day's activity are small in total volume compared to the file itself, can be handled by a much smaller erasable file which forms part of the total file system. Each night, the entries on the erasable file can be transferred to the permanent photographic file.

Of course, the permanency of photographic media also implies that sorting, collation, and interfiling are not easy. Meeting this objection is the basic design consideration underlying Eastman Kodak's so-called Minicard system, with which you are familiar. It is my opinion that there is a better answer, and discussing it will lead us back to a discussion of the interaction of methods and machinery which is the main purpose of this paper.

In the case of human records, as opposed to the machine records we have been talking about, file management usually involves:

1) Ordering the record. Entries are filed alphabetically, or on some other systematic scheme, and new entries are interfiled in ther proper order relative to those already filed.
2) Purging the file. When a new entry replaces an old one either partially or completely, the old item is removed and discarded.

[2] A. W. Tyler, "Annals of the Computation Laboratory of Harvard University," vol. 16, pp. 146–150; 1948.

[3] G. W. King, G. W. Brown, and L. N. Ridenour, "Photographic techniques for information storage," Proc. IRE, vol. 41, pp. 1421–1428; October, 1953.

3) Providing random access to any item of the stored information.

These requirements are so familiar that they seem to be inescapable. Actually, they are requirements which have been imposed on us by the slowness of the reading process. Random access to any item of the record is required only because we cannot read through the whole record rapidly enough to find what we want, no matter where it is located. Ordering the file is a help to random access; it lets us know roughly where to look for a stored item. Purging the file is required to keep the total size of the file down to manageable limits, and is otherwise necessary because we do not wish to read the file exhaustively in order to determine which items of information are modern and which obsolete.

So far as files kept and consulted by human beings are concerned, these are all sensible restrictions. But when we talk about files that are to be kept and consulted by machines, we must question them. In magnetic-recording technology, a single reading station can handle about 10^5 bits per second; the reading station of a photographic file is between one and two orders of magnitude faster. Let us make the conservative assumption that we have the ability to read the file at a rate of 10^6 bits per second per reading station. Given a hundred reading stations, then, we can read a file containing 10^{10} bits in 100 seconds, or less than two minutes. In the course of this exhaustive reading of the file, we can find all entries that are wanted, whether the file is ordered or not. Having found all entries, we can readily determine which of two or more contradictory entries is the more recent, and therefore the one to be believed. We can conduct many searches in parallel while reading through the file, so that the effective random-access time can be quite low, on the average, even though we have made no formal provision for random access.

Thus we see that a machine file—because we can read it exhaustively in so short a time—does not require to be ordered, does not require to be purged, and can provide random access to any item of the information it holds without the elaborate provisions for random access which are so familiar in the case of human files. High reading speed is a substitute for sorting, for interfiling, for purging, and indeed for most of the things that are commonly thought to be inescapable when a filing problem is considered.

These remarks have only an heuristic value, and indeed only an heuristic intent. The development of large machine-memory systems will go on its way despite them. But it does seem important to realize that the filing methods we are all familiar with are inappropriate to the capabilities and requirements of the information-processing machines we are now learning how to build.

Discussion

C. Mooers (Zator Co.): We are all familiar with the many difficulties in terminology in the computer and related fields. This paper seems to be leading to another such confusion. "Information Retrieval" has been in use for five years in connection with machines for discovering information in a library of documents or books. A simple computer memory system does not of itself perform "Information Retrieval." In the sense now quite well established there must in addition be some sort of scanning and selection upon the record in the memory, giving an output of a document call number. Thus "Information Retrieval" should be distinguished from the simpler "Memory Access."

Mr. Ridenour: This is an entirely valid comment on what I actually said. At the time I submitted a title for this paper, I had intended to go on to show how the ability to perform rapid and complete reading of a large file eliminates entirely the necessity for the indexing which is the fundamental basis of information retrieval in the sense that Mr. Mooers uses the term. As it turned out, I omitted this from my discussion, since it would have required more time than was available. I apologize to Mr. Mooers for not discussing the problem which is closest to his heart.

R. W. Hughes (Federal Telecom Labs.): Please comment on magnetostrictive lines for computer memory use.

Mr. Ridenour: To be honest, I have had no experience with such lines. I believe that they are quite comparable with other types of acoustic delay line in having the same virtues, such as high reliability, and the same disadvantages, such as relatively high cost per stored digit.

J. L. Hill (Sperry Rand): Do you have any offerings of logical or arithmetic systems commensurate with the proposed 10 mc reading rate you indicate substitutes for interfiling, purging and random access?

Mr. Ridenour: If the circuits external to a memory system cannot work at the reading speed, the combination can usually be brought back into balance by providing several parallel logical circuits which work out of a single reading station, which latter operates serially by definition.

J. C. Hosken (Farrington): How do you put the information into the photographic memory?

Mr. Ridenour: At Eastman Kodak this is done by means of a system involving a punched-tape reader which operates a series of shutters in front of lights which then either do or do not expose a corresponding square area on the photographic card. While there are other ways of entering information, this is a perfectly straightforward and satisfactory one. The basic requirement is to write into the photographic medium a great many bits at a time. When this is done, the writing rate can be very high. In this connection, it is interesting to observe that one 200-foot reel of 7-inch aerial reconnaissance film transported 1,000 miles in a B-47 aircraft amounts to a communication channel that is about 10^8 bits/sec. wide.

B. Lippel (Signal Corps): Do you think there would be merit to a large permanent file—for example, photographic—paralleled by a smaller magnetic file with additions, deletions, and so forth, arranged so that the latter can override the former if it holds the entry wanted?

Mr. Ridenour: Yes, indeed. You just invented how to do it. It has to be done that way.

R. E. Wright (Teleregister Corp.): Do you know of any work being done on a large random access memory utilizing physically small magnetic cards?

Mr. Ridenour: No, I don't. I have heard talk here and there but I don't know of anyone working on that.

Moderator Weiner: The following comment from John F. Rafferty of the National Bureau of Standards is addressed to Professor Adams. At the SEAC computer installation at the Bureau of Standards the programmers were interested in devising a way to record the engineers on magnetic tape, read them into the machine and press the "memory clear" button. Obviously, as an engineer, I would be interested in progress reports on the attempt to program the programmers out of business.

The Role of Communications Networks in Digital Data Systems

R. C. MATLACK†

Summary—**Communications have played an important role in linking together teletypewriters in manual and automatically switched networks. There is a rapidly growing need for higher speed communications to transmit digital data from remote to centralized locations. New engineering and maintenance approaches are necessary to take advantage of the existing widespread voice communication facilities. Since this plant has been primarily designed for voice transmission, new techniques are desirable to insure reliable high speed data transmission and switching. This would be facilitated by use of a universal language and standardized speeds of transmission. The speaker will discuss these problems as viewed by a communications supplier and will present his views on future development.**

I AM SURE that we have all dreamed of the future possibility of having quick and easy access to an electronic computer in order to obtain from its memory a fast and painless answer to a question. As extensive decentralization has taken place in industry, it has become more difficult to keep top management informed with *up to the minute* data regarding all phases of operation. It is almost universally believed that this difficulty will be overcome by high speed transmission of current information to a centralized data processing center from a large number of remote locations. Obtaining easy access to a computer from remote locations and keeping its memory current, is a problem of communications requiring the use of a digital language. In recent years, even the general public has become accustomed to seeing information represented in coded form on punched cards. In the Bell System, we have developed extensive record communication services based on punched paper tape.

Coding by marking or punching is an old art. The speed capability of electronics has necessitated the development of new recording methods such as magnetic tape which, in turn, accentuates the need of higher speed communications systems. Computer engineers have suggested that the existing widespread telephone network may be utilized to rapidly accelerate the use of systems which will automatically process practically all types of existing business records. Before discussing these future development possibilities, I would like to briefly discuss existing data communications networks in order to provide some background information which I feel will help in an understanding of the future, since most developments are evolutionary rather than revolutionary in nature.

Communications have played a vital role in the transmission of data ever since Morse devised his well-known code in 1838. This was followed shortly thereafter by successful telegraph printers. In 1874, Emile Baudot greatly accelerated this progress with his 5 unit selecting code which became, and still remains, the common language code for telegraph printers, more familiarly known as teletypewriters.

Two general types of services are now furnished by the Bell System. These are known as (1) teletypewriter exchange service, commonly referred to as TWX and (2) private line services.

The Bell System TWX service is furnished to fill the need of customers who have a limited amount of point-to-point data to transmit per day or have the need to communicate with any of a large number of other TWX customers. An example of this latter case is the need of a travel bureau of a company to quickly arrange and confirm wide-spread hotel reservations for its employees. At the present time in the Bell System, there are approximately 37,000 teletypewriter exchange stations employing 3,352,000 miles of circuits. The system averages approximately 89,000 calls per business day with each call averaging 8 minutes. Messages are transmitted either *on line*, as they are typed by the keyboard operator, or automatically, by placing a previously prepared perforated tape into a transmitter distributor. Speed is limited in the *on line* case to the capability of the operator whereas the automatic method is limited by the system capability, which is 60 words per minute.

The second, or private line type of services are quite varied in nature. There are service offerings of teletypewriter, voice, telephoto, facsimile, telemetering, program, and television. In the data transmission field, these private line circuits are capable of information rates varying between the limits imposed by a 170 cycle telegraph channel to a 4.5 mc television channel. At the present time, the largest Bell System private line service is teletypewriter. It runs the gamut from leasing a pair of wires to a customer who furnishes his own teletypewriter terminal equipment, to the leasing of a complex automatic switching system and wire network along with the associated teletypewriter equipment. These large automatic systems compare in functional complexity to modern computers. They play a vital role in data processing systems now in use by many large companies.

Private line teletypewriter service is furnished to a particular customer for his exclusive use for a specified period. These periods range from eight hours per day on specified days, up to twenty-four hours per day, seven days a week. A few typical arrangements are: 1) point-to-point service, 2) party line service, where a

† Special Systems Engineer, Bell Telephone Labs., New York, N. Y.

number of stations are all connected to one line and one or more are arranged to only receive; this type of service is arranged for various press associations who sometimes have as many as 300 to 400 stations on a line, and 3) large automatic message switching systems. A typical large system in use consists of 2 switching centers, 94 sending stations, 143 receiving stations, and 23,000 circuit miles. Daily traffic averages 48 thousand messages with an average length of 17 words. Automatic switching is accomplished by the use of address characters preceding the message and by end of message characters following the message. Bell System private line teletypewriter services now aggregate approximately 42,000 stations employing about 4,000,000 miles of circuits, which is slightly greater than the TWX service with regard to both stations and circuit miles. A cost of $0.11 per message is experienced by one company which has an efficiently loaded, large system. In this case, the message averages 40 words in length and travels 500 miles. This cost represents only the revenue paid to the Bell System and does not include additional company operating costs.

Teletypewriter services now in use have nominal speed capabilities of 60, 75 or 100 words per minute, with corresponding bit rates per second of 46, 57, and 74 based on the standard 5 unit code. An average teletype word is assumed to be 5 characters plus a space. No error detection facilities are provided, but these services are engineered to rigid transmission standards. Circuits are engineered from end-to-end so that the error rate is never worse than one in 44,000 characters. Average telegraph circuits, however, have transmission error rates of about one in 8,000,000 characters. These error rates include the normal random summation of various sources of transmission distortion, but do not include distortion due to interruptions, static, magnetic storms, etc. The above error rates cover only the transmission and do not include operator amd machine errors. A regenerative repeater which retimes the signal elements is inserted in over-all circuits comprised of many links, in order that the transmission standards will be met.

The teletypewriter services were planned primarily as a means of communicating ordinary text. Because teletypewriter machines do not use an error detecting code, computer people have sometimes been reluctant to use the service for business and other nonredundant data. In view of the high degree of error control maintained, as described above, the teletype service is suited to many business uses. The automatically switched networks should, in particular, be of interest to those designing integrated data processing systems. Work is underway to develop an economical and reliable error detecting scheme for use with the 5 unit code.

Having looked at the present facilities for transmitting data, let's look at the future. In the first place, it is now clear that high speed digital data communication is going to be necessary to support the further mechanization of business and industry. At the present time, there is a growing need to transmit data in the fastest possible manner from widespread remote locations to centralized data processing centers in order that information there may be as current as possible. In fact, we in the Bell System have had numerous inquiries from manufacturers of computers and business machine equipment, as well as from users of such equipment, as to what developments are planned for use over private line networks and for use over the telephone network.

In the private line field, it is possible to utilize channels which vary in discrete steps of bandwidth all the way from telegraph to television. Examples are: voice channels with upper usable frequencies varying from 2,400 to 3,300 cycles; 5,000 and 15,000 cycle program channels; carrier group channels such as 48 and 96 kc; and 4.5 mc television channels. At the present time, only the voice and telegraph channels are being used for data transmission, but the broader channels, along with suitably designed modulators and demodulators will be considered when the demand materializes for higher speed data transmission systems. We have, under development for the SAGE project, a data transmission system capable of transmitting 1,600 bits per second over a private line facility. In order to realize this speed of transmission, the circuits must have some delay equalization treatment and some of the carrier systems must be adjusted to improved impulse noise standards. Also under development, is a double sideband amplitude modulation system with an expected capability of transmitting 750 bits per second over a standard voice bandwidth private line facility. The Teletype Corporation has, under development, a punched paper tape transmitter and receiver which can be arranged for 5, 6, or 7 bit codes and a speed capability of 60 to 100 characters per second.

Computer engineers have made numerous inquiries regarding the possibility of the Bell System's using the regular telephone switched network for data transmission. It has been pointed out that this is particularly attractive to small businesses and agencies who have limited amounts of traffic and desire to transmit small volumes of traffic on an intermittent basis. Since this network is practically universally available and is rapidly being arranged for customer direct distance dialing within the United States, it has very attractive possibilities for this use. It has been proposed that the telephone be used in the usual manner to establish a connection to a data processing center. As soon as the connection is established, a suitable data input device would be substituted for the telephone.

I would like to discuss now some of the engineering problems in connection with this particular proposal. As you know, the telephone plant has been basically engineered for voice transmission and the only criterion is that a satisfactory voice transmission circuit will be furnished. However, voice transmission requirements are much less severe than data transmission requirements, particularly with respect to impulse noise and delay distortion. The telephone plant is made up of many different kinds of voice and carrier facilities which vary widely in bandwidth, delay, and noise characteristics. Slow speed data transmission in the form of direct current dial pulses and numerous alternating current signaling and data transmission schemes are widely used to establish and disconnect telephone connections. Several years ago, the Bell Telephone Laboratories initiated a research program to acquire fundamental knowledge of high speed data transmission over telephone circuits. The main purposes of this program were to explore faster and more efficient ways of establishing telephone connections. Much has been learned from this and other similar projects regarding high speed data transmission over telephone circuits.

Extremely high accuracy of data transmission is a must, and this presents a most pressing problem. The general trend in the computer field has been to add parity checking to all data processing equipment and arrange the characters in block form on magnetic tape with space between the blocks. Block lengths vary between approximately 70 to 120 characters. A likely way of transmitting them is by start-stop operation, whereby checking equipment at the receiving end will, at the end of a block, send a return signal to the transmitting end to advance to the next block if there were no errors, or back up and repeat the block in the event there were one or more failures to parity check or check that the correct number of characters per block were received. A considerable amount of study and experimenting is necessary to determine the optimum block length for transmission purposes. If the block is too long, too much time will be wasted by retransmission of blocks that have only one error. If, however, the block is too short, the return signal time will become an appreciable part of the transmission time. This is greatly aggravated by echo suppressors which are used on electrically long circuits in order to prevent the annoying return signal. Transmission is possible in only one direction at a time since the opposite direction is purposely suppressed in order to avoid echoes. As the direction of transmission is reversed, the echo suppressors do not immediately reverse, but take as much as 120 milliseconds on a long circuit such as from coast to coast. This is not noticeable in voice transmission, but is a time loss with data transmission.

Many of our modern carrier systems are designed with compandors which is an abbreviation for compression and expansion. Their purpose is to compress loud speech and expand weak speech at the transmitting end and reverse this process at the receiving end. By means of these compandors, very decided signal-to-noise advantages are obtained for voice, and transmission is much improved over noisy circuits. These decided signal-to-noise advantages lead, of course, to an engineering of the telephone channel with a correspondingly greater physical noise before the compandor is inserted, than would be the case if the channel were to be used without compandor. This higher noise level is detrimental to pulse type transmission and therefore special design measures are necessary to make a high speed data transmission system function reliably over compandored facilities. In spite of these difficulties from a technical viewpoint, it appears possible to design a fast, highly reliable data transmission system suitable for use over the telephone switched network. The primary function of the network, however, is to carry telephone conversations and this function must not be jeopardized. Special input and output lines (*i.e.*, subscribers' loops) and additional facilities to increase the traffic capacity of the network probably would be minimum requirements.

It is interesting to speculate as to what type of services might be provided. Indications so far are that there will be need for two distinct types: a very simple slow speed arrangement whereby a message can be sent by a simple keyset to a computer or machine and a return answer by voice from a computer or machine; and a means of sending all forms of digital data from tape to tape at the highest possible speed. An example of the first case would be if you wanted to find out how many piece parts of a specified number were in stock at a given factory. Connection to a computer might be achieved by dialing a telephone number type code, followed by a proper code to determine how many of the desired piece parts are available. The computer would give the answer in digital form, or by means of a convertor the answer might be transmitted back to you in voice. This is, of course, a very simple arrangement and many variations of increasing complexity are possible. For example, the question could be transmitted by means of a teletypewriter and the computer would type the answer on the teletypewriter. This type of service is *on line* transmission, but since it is visualized that a high percentage of this type of traffic would be local in nature, high speed is not a vital factor.

A large number of speculative types of service are possible which involve the second type of service which was previously mentioned. These types of service involve the transmission of all forms of digital data from tape to tape at the highest possible speed. A typical example would be a typewriter keyboard associated with the necessary equipment to code the characters on magnetic tape as they are being typed. When desirable, the tape would be placed in a transmitter and the information sent over a connection after being established by dialing in the conventional manner. The necessary equipment at the receiving end would place the signals on a tape which would be a duplicate of the sending tape. Included would be the necessary checking equipment to advance the transmitting tape block by block as previously discussed. A great many variations of this basic plan are possible. For example, one might conceive a mixture of all different types of digital data placed on the tape in message form such as computer data, punched card data, and teletypewriter data. This would require that messages of different types be preceded by characters indicating the type of data and followed by end of message characters. It would be possible to place the tape in a suitable reader associated with switching equipment which could monitor the code directing characters and transmit the message to the proper end equipment such as computer, card reader, or teletypewriter. It is obvious that this type of switching requires a common language code in order that certain characters can be reserved for only that use. I would like to urge that serious consideration be given to standardization in the field of data transmission. I realize that rigid standardization is undesirable because it tends to stifle the progress of a newborn art. I feel, however, that some reasonable amount of standardization which will not block progress is possible and will avoid a lot of future headaches.

I would like to emphasize that the discussion regarding the use of the telephone switched network has been entirely speculative in nature for the purpose of stimulating some thinking regarding the long range aspects of the data transmission field. I would also like to acknowledge that much of the information regarding the existing TWX and private line systems was obtained from Messrs. L. A. Gardner and E. R. Robinson of Bell Telephone Laboratories.

Concluding ideas are summarized as follows:

1) The continued mechanization of business will require expanded digital communication facilities. Unless the communication problems are solved, the full benefits of mechanization will not be realized.

2) The computer and the communication industries each need the cooperation of one another in working out standards for data transmission.

3) And finally, let me say to all those designing integrated data processing systems, "Don't leave the communications problems until the end. Call in a communications man early in the project and use his specialized knowledge to help you design an optimum system."

Discussion

J. H. Waite (R.C.A.): Are you aware of any efforts being made to develop error detection and auto-correction techniques for teletype communications?

Mr. Matlack: Yes, as mentioned yesterday by Mr. Fitzgerald, although I am not aware of it, Western Union is working on that and I mentioned that the Bell System is also working on it. Of course, an obvious way of including error detection is to transmit the message twice. This is a waste of time and better methods can be used. Error correction requires much more redundancy and we feel it is sufficient to be able to know that there is an error without correcting it; that is, action can be taken if you know an error is there rather than include all the redundancy necessary to correct the error.

M. S. Rubin (Filene's): Is it now possible to transmit seven channel magnetic tape data over a conventional telephone wire, not teletype? If not, what can you tell us about current work being done?

Mr. Matlack: It is not presently possible and, as I indicated to you in my conclusion, I have only spoken about the engineering considerations of this. There are many policy questions involved which have not as yet been resolved regarding the use of the telephone switched network for this type of service.

Capt. B. R. Jacobellis (Signal Corps): What form will the error detecting code take?

Mr. Matlack: Is this with respect to teletype or the future type system? In the case of the teletype system there are several possible ways of doing this, such as on a per line basis. When the carriage returns to the next line it is possible to add across the rows of characters and come out with what is equivalent to a sum, odd or even, and type whatever random character happens to come up there. That is one way of doing it and was mentioned by Mr. Fitzgerald yesterday.

There are other ways of adding every so often, say every 20 characters, or 30, or whatever you want to pick and how accurate you want to make the system. You can add a redundancy check character and then remove it at the terminating end.

With regard to future systems I think it leans towards a seven bit code with one of these as a check bit.

E. L. Lach (Western Union): What percentage of accuracy can be expected with the five channel error check under development with your firm?

Mr. Matlack: I can not answer that at the present time because we have not tested it.

C. K. Budd (Signal Corps): Is the Bell System interested in standardization of digital code language and if so, what is the present status of work on such a project?

Mr. Matlack: Preliminary discussion with interested computer manufacturers and users has been started under the sponsorship of the JCC. I think this will come out more in the panel discussion that is coming along. I would prefer to delay this question until that time.

W. L. Poland (Daystrom Instr.): In transmission at 750 bits/sec over typical voice lines, what is the expected average error rate?

Mr. Matlack: This is going to depend, as I mentioned in the paper, on how long we make the block length. Nothing has been picked as yet and tests are just now beginning to determine error rates. It is very dependent on how long the block length is. If you make the block indefinitely long there is bound to be an error in transmission to upset some of the characters. However, if the block is only a few characters long the chance of an error is practically impossible. Therefore, some compromise has to be effected as to block length.

H. E. Tompkins (Burroughs): How does the ratio of cost to bandwidth vary for the different bandwidth channels available?

Mr. Matlack: In the private line field in the Bell System the higher you go in bandwidth the cheaper it is in terms of bits. For example, consider slow speed teletypewriter at 100 words a minute. The private line charge is roughly $1.50 a mile per month. An ordinary voice channel, which has, instead of 100 words per minute, a capability of roughly a thousand words a minute, is $3 a mile per month. That goes up to a full 4.5 mc which is in the order of $35 a month. There are additional terminal charges for connecting it in at the time it is connected. However, that gives you an idea. The cost per bit is going down tremendously as bandwidth goes up.

J. McDonnell (IBM): A large computer installation costs $2 to $5 per hour to operate. Do you know the comparable costs for voice, radio, carrier and TV line for a 1,000 mile station to station operation?

Mr. Matlack: No, I can't give you the comparative breakdown of that. Of course, our voice facilities are carved out of several different things, such as our coaxial cable and our microwave relay. All of those are in use for providing transmission. I think you do know pretty well, if you take a station to station toll charge rate on the telephone, a thousand mile call would be roughly for three minutes on the order of $1 to $1.50. That would be my guess without actually looking at a rate schedule. In this three minutes you could transmit in the order of say at the rate of 700 bits per second and that extended over your three minute period you are pumping in a fair amount of information, particularly if it is coming from one isolated spot, and this is the point we should all keep in mind.

When you eventually get down to it someone has to make these records. When you talk about millions of bits you are talking about a lot of people originating data. In cases of short transmission time you can handle the output work of small offices quite easily. When you get into the higher echelons with computing centers, the need for high speed transmission and longer times becomes necessary. Where you can take a private line facility and keep it loaded a high percentage of the 24 hours it is the most economical method.

E. Donahue (IBM): Has Bell experimented in the field of data transmission through short wave radio hookups and, if so, what results have been obtained?

Mr. Matlack: No, we have not expressly experimented with the short wave radio. However, our ordinary teletypewriter transmission facilities are going over microwave relay in some cases from coast to coast. There are voice channels carved out of the microwave system which are carrying teletypewriter messages.

L. J. Score (Commercial Controls Corp.): Has any thought been given to six, seven, eight bit transmission with parity or other checks?

Mr. Matlack: Yes. As I mentioned, I think, in connection with an earlier question on the longer range aspects of this problem, we are considering a seven bit code. Our difficulty with the teletypewriter five unit code—and this is in connection with our automatic switching system—is that we must avoid a code which will cause disconnection of that system. Therefore, you cannot put over existing systems a new code unless that code is controlled and, of course, requires new equipment design for the code. So it is an economic question and we feel that in order to use the equipment the thing to do, where high accuracy is a must, is to provide some other type of parity check within the five unit code. As we go to newer high speed facilities our thinking is that it will be in the direction of the seven bit code.

W. A. Hunt (IBM): How often, on the average, do noise bursts occur and how long do they last on lines such as the SAGE type?

Mr. Matlack: This is practically impossible to answer. These noise bursts occur in a random fashion. For example, lightning comes along and you get a burst of high impulse noise in the line. Some switchman working on the frame accidently touches a tool against the line causing a burst. Some error rate will be specified as the result of tests underway. This will include all factors causing errors such as impulse noise.

I mentioned in the teletypewriter field we are engineering in the worst case to an error of one in 44,000 characters. As you increase the bandwidth, you obviously let in more noise. However, depending on the measures you take, you may still be able to get a very high error-free rate so far as random noise is concerned. But it is such a random occurring thing that there is no real way to say how often it will occur.

Standardization of Computer Intercommunication

H. R. J. GROSCH†

I AM GOING to start today by taking a little different approach to the problem of standardization. Mr. Matlack and members of the panel are concerned primarily with the specific problem of standard media, but I would like to speak, at least for a few minutes, about the philosophy of standardization in general rather than about our particular problem. What is standardization? What are the permissive functions of the American Standards Association? What, indeed, does a printed standard of the ASA do for its sponsors and for newcomers in the technical field?

Many times, in discussing this area with individuals and with organizations, I find there is considerable ignorance and even fear about the forces behind the standardization movement. I can't speak to you as an expert, although I have been on several ASA subcommittees, and I am not authorized to speak about broad policy, but you should know something about the philosophy of the standardization movement. You should know, for instance, that standardization is a means of registering good methods of achieving technical ends. Decisions in the standards committees are made by consensus. If a consensus cannot be achieved no standard is issued. Standardization is merely a recognition of this consensus. No member of the standards committee involved, be he an individual member or a society member, is obligated to follow the adopted standard in any way. There is no such obligation implied or expressed in the philosophy of standardization or in the constitution of the American Standards Association. On the contrary, mechanisms are provided to register standard designs, standard methods, so they are available for those who wish to use them.

Senior people in our field, such as Bell and IBM, often hold back because their experience is so great they don't want to have their activities restricted by the intrusion of novices. There is no chance of that happening—I repeat, *no* chance. If someone wants to invent a 13-hole paper tape they have the express permission of the ASA to do so. The dimensions and tolerances of that tape should be registered so if someone else wants to have a 13-hole tape he can find out how his predecessor did it, but there will be no pressure to abandon the 13 for two 7's.

I want to talk for a few minutes about the computer art and the way it has developed. People are talking about intercommunication between million dollar machines over thousands of miles of distance. Yet eight years ago a prime article of commerce was a punched card table of reciprocals. And with this tremendous growth has come a desire to exchange a great deal more information than was necessary in the early days. We used to trade IBM panel diagrams and things like that.

† Manager, Investigations; Aircraft Gas Turbine Development Dept., General Electric Co., Cincinnati, Ohio.

But now there is a present and pressing need for the trading of mass information.

In fact, I would like to point out that Mr. Matlack is being very conservative when he talks about voice transmission and the Bell System. It is my belief that within a few years transmission of digital information is going to overwhelm the transmission of crude human voices. We are participating in a technological burst which is going to alter such things a great deal.

As the mass flow of information accelerates it is more and more important that we make a clear statement of the problems which warrant standardization and set up some kind of cooperation to tackle these problems. I think of three areas where the transmission of masses of information in business—remember, this is supposed to be one of the themes of our conference—is becoming increasingly important. First is the centralized company where information has to flow radially between the outlying sales offices and factories and the central organization. True, many of you have heard me say that there are no human beings at the center capable of listening to this information, but if we replace obsolete executives with large computing machines this becomes a practical method of running a business.

A second class of transmission of large quantities of data arises from a decentralized operation like General Electric. Here there is a flow over most of the meshes of a rather complex network. There is no single point where all information must concentrate. Requirements of band width and tape length will be different. Nevertheless, the total number of bits transmitted per day is undoubtedly much greater than in a radial system. Moreover, it is not out of the question that an outfit that wishes to decentralize authority—and that is fundamental in G.E.—may still wish to centralize information as to how that authority is being used or misused.

There is still a third and in some respects more interesting possibility than the other two. That is the serial communication of information up or down between supplier, the main business, and customer—a really enthralling possibility. We have had a slight touch of this in my operations in General Electric, where we have sent information about our new engines to computers in aircraft companies on the West Coast. We have sent the information in computer language; not, it is true, over telephone lines—the time schedule was not that demanding—but at any rate the 701's talked to each other in their own language without too much interference from human beings. We regard ourselves as the "business" and North American, say, as the "customer." This can be inverted and North American might regard us as the supplier and send the information on the combined engine and airframe on to the Air Force as a customer. The Air Force might submit it in still further automatic transmission form to "Aviation Week"!

Here are at least three imaginary situations; each of them different, as regards the way in which communications travel, and each of them capable of generating tremendous amounts of information. It is unnecessary to list all the various channels and physical media that may be involved in this intercommunication, ranging from the old punched card to the fanciest multichannel magnetic tape.

How many people in this audience know that there is no formal American standard for paper tape? There is no standard issued by the ASA or any international association that gives the dimensions and tolerances and paper stock requirements on even five-channel tape, let alone the less common types. I was shocked to hear this from ASA, especially when I found there was already a standard on remote and telemetered supervisory equipment, pumping stations on pipe lines and so on, for which one can pick up a telephone and dial a number and get a meter reading miles away. There is a standard on something that specialized, but not on paper tape!

The hardware involved in computer intercommunication may be divided into three classifications. First, and most important, is the primary hardware which is capable of recognizing a signal or generating a signal. Here I am referring to the type of magnetic recording, and to the lateral spacing and number of channels across a magnetic tape. In the case of perforated tape, the equivalent properties are the size and position of the holes, and thickness of stock. In punched cards, IBM 80-column, Minicard, Remington Rand, and other types each have fixed dimensions and tolerances, and material requirements. If this single most important item is standardized it is then possible for various input and output mechanisms to at least generate and receive messages.

Second, there is a hardware problem involved in the *acceptability* of the signal, In many magnetic and paper tape systems, there is a built-in redundancy check which makes certain combinations of signals which get by the head of the tape unit, unacceptable to the rest of the electronics. One might say, standardization of primary hardware is required if the computer is to hear a signal, and of secondary hardware if it is to pay attention!

Finally, if the message has gotten past the hardware, there is the third element of agreed code. A general purpose computer is perfectly willing to accept, to coin a phrase, standard gibberish because it is capable of translating that gibberish into its own language by programmed manipulation. So, if the bit code internal to a computing machine uses a certain set of ones and zeros to represent the letter "f," and if that set happens to be Western Union notation for the letter "a," no harm has been done providing a general purpose program is used as part of the communications network to interchange "a" and "f" in the message.

The first item of business, then, is insurance that the message can be gotten into the system at all. We should talk about tape punching mechanisms, numbers of channels, and recording heads. Secondly, we should insure that the signals are acceptable to the logic of the communications system. Lastly, if possible, we should register common codes so that "a" will always be "a" and "7" will always be "7," or so that we can program the translation.

So much for the individual characters. Now, how about the problems of semantics, the problem of doing something useful with this set of bits which was sent from computer to computer. This involves the user of the machine as well as the builder of the machine. I think many of this audience are familiar with efforts on the West Coast to standardize computer language in the broad sense of the word. First, something about motivation: many of the users of digital systems have been and will be faced with the problem of changing machines. During 1956, for instance, my group will be operating a 701 through at least April and will be beginning to use three 704's. These machines have different languages. They have rather similar character representation and identical reading equipment. But programs we have written for the 701 are not usable on the 704. And how much more difficult the situation would be if we were switching from a small IBM machine to a larger machine built by a competitor with very different philosophy.

One way to minimize the effect of a changeover is to have written your programs initially in a pseudolanguage—a pseudolanguage which both the 701 and 704, say, can translate into their own preferred language. This means that if the 701 programs were originally written in a pseudolanguage, which compiling and interpreting routines then translated into "pure 701," a different set of routines could be used to take original programs in the pseudolanguage and translate into "pure 704" with no further thought on our part. This is a very long term problem in standardization, a very difficult one, but one which may properly be given as a charge to some working committee.

It is clear we face a dynamic situation, not a static one. Grace Hopper has been talking about this for years and those who have heard her speech eight or ten times have observed trends; there are changes. There are going to be further changes, especially among those of us just now getting on the bandwagon. We must regard standardization of hardware and semantics as temporary. And we have already passed the point of optimum beginning. Our practices are already divergent in many respects. I am not trying to call for restrictive measures. Someone pushing 31-channel tape? Fine, by all means let's try it, but let's not have two different *seven*-channel systems if we can register a permissive standard that all new designers can refer to. Usually the saving in design cost to the late comer will be so great that he will willingly—I repeat, willingly—conform to the established standard.

This will help hundreds of users of large machines and thousands of users of small machines in the next years. Our industry was worth a few thousands a month ten years ago and will be worth a billion dollars a year by 1962. It is important that we move to help ourselves by using these techniques of standardization at the earliest possible moment through member societies of the JCC, in the JCC itself, in the fraternity of computer users, and through the manufacturers.

Discussion

C. Adams (Westinghouse): Would you care to pick the date on which the message volume will exceed that of TV and that of voice?

Mr. Grosch: I don't care to, but I will.

Mr. Adams: When will this information volume exceed the volume of bits transmitted by the United States mail?

Mr. Grosch: I will make a guess at the first one. As we improve the technique of TV transmission, perhaps by sending only the derivative instead of the complete image, and as integrated data processing in digital form goes up, this intersection might be as soon as four years from now.

As for the second question, one school generation, at the present rate!

M. Harper (Sperry-Rand): How efficient will your translated 701 routines be when run on the 704?

Mr. Grosch: We decided that the efficiency would be so low that it would be better to use the pseudolanguage idea and rewrite some of our 701 programs in the pseudolanguage. It is difficult to think of a method of bringing in the indexing capabilities of the 704 when translating a 701 program. Of course, the opposite would be more easily possible; one could call in for each indexing instruction a complete subroutine of conventional 701 variety; but to go the other way requires the recognition of a pretty large chunk of program.

H. Livingston (Burroughs): What rules would be used in establishing a pseudolanguage? Would the pseudolanguage apply only to the foreseeable future?

Mr. Grosch: Part of the problem is certainly caused by the unforeseeable extra capabilities of future machines; part, by the somewhat unforeseeable utilization of these machines. Look, maybe it isn't as desperate as all that. We are talking quite a bit at this meeting about business problems. The language of business is messy: it is historical, it is poorly chosen, it is extraordinarily redundant. But the total dictionary of business terms—and I mean information, not just words—is really pretty limited. I would think, for instance, that it would be much easier to generate a pseudolanguage for accounting problems than differential equations. Most of you know that there is a strong effort to generate a pseudolanguage in which the input would be algebraic, and I just can't believe that the problems of accounting require a more imaginative language than that. What those problems require is more imaginative *accountants.*

R. Gillett (NSA): Does not the lack of standardization mean that the investment in programs makes a user reluctant to change to a new computer?

Mr. Grosch: Well, as long as there is a shortage of new computers I don't suppose this is a governing item. In talking to my management about moving to bigger and better and more expensive machines I usually say that the way to get around the problem is to have a highly professional group which in the first place is willing to learn a new language, in the second place is capable of cooperating on a pseudolanguage, and in the third place, is able to change over efficiently. I suppose if the day ever comes when there are surpluses of large computing machines, and price wars between the manufacturers—that will be the day!—the quality of programming of these conventions by the manufacturers will be the deciding item in the switchover.

Mr. Schleicher: Has a good dictionary of computer terminology been compiled and, if so, where?

Mr. Grosch: There have been several glossaries. The IRE Committee on Electronic Computers had a subcommittee on nomenclature of which I was a member. Grace Hopper helped work up another. There are two or three such efforts published—one by the IRE, and one by Ed Berkeley. Again, I think the crux has been the feeling that it was restrictive. A lot of people felt that once there was a definition of "flipflop" or "programming" they were stuck. That isn't the purpose. A standard on terminology simply warns an author that he should explicitly define new words, or words he uses in a non-standard way. It would be perfectly proper to have a subcommittee of the standardization effort I am describing consolidate these efforts.

I wonder if I might mention again the word "consensus." No standard is published by the ASA unless a consensus of the members of the working committee has been reached.

Statements from Manufacturers on Standardization of Magnetic Tape Records

M. E. Femmer (IBM): Ladies and gentlemen, IBM is very pleased to be invited to participate in this discussion. We hope also to be able to participate in future programs of this type. Perhaps of more significance, we will welcome opportunities to work with our associates on committees and similar groups to help establish suitable industry-wide magnetic tape standards.

As an engineering and manufacturing corporation, we have long recognized the suitability and desirability of standards in all phases of our operation. We have used them internally to simplify design, reduce costs, and provide greater manufacturing flexibility. We have on many occasions worked with others in establishing standards in areas of activity outside our own organization. Our past experiences encourage us to continue in this type of activity.

The Data Processing Field is an extremely vigorous and dynamic one. It has grown in a very few years from scattered laboratory efforts to an industry in its own right. Indications of continued growth of the Computer Industry make present consideration of standards very timely. Magnetic computer tapes are even more recent than the processing systems. It is not surprising, then, with such growth, that so few standards, if any, have been established in the tape area. As working participants in this active field, we in IBM feel that standardization should be approached with real intent but with some caution. It must always be clear that the standards adopted will not stifle or impede rapid technological progress which is certain to continue for some time.

In this context a standard might be loosely defined as an agreed upon common specification in a particular area. If the area of activity shifts rapidly as a result of a significant technical break-through, the present common specification is no longer applicable, and additional standards become necessary.

It should be further noted that we are not attempting to standardize on particular designs. We do, however, endorse standardization that is meaningful to the customer or user of the equipment.

We have a genuine interest in the success of this program and with such an interest we would like to suggest a starting program of sufficiently large scope that it will test our mutual abilities to work together on an industry-wide basis. The starting program should not, on the other hand, involve so many considerations that it might prejudice the early success of the venture.

Specifically, as a starting program, we suggest that an attempt be made to standardize on the physical characteristics of tape. This would include such characteristics as dimensions, tolerances, weight per unit length, and tensile strength. These are basic characteristics but extremely important. Without agreement on these, other standards will have little value.

In addition, we further suggest, as a starting program, that an attempt be made to standardize on tape reels from an engineering point of view.

These will be significant first steps. As a minimum, they will insure that tape is completely physically transferable from one manufacturer's tape unit to another. If such a starting program can be successful, we encourage the further consideration of additional standardization measures.

Again, may we express our genuine and continuing interest in this program, and also our thanks for this opportunity to express our position.

R. B. Lawrence (Datamatic Corp.): Datamatic Corporation is happy to be invited to participate in this panel. I would like to call to your attention a recent document which some of you may have heard of. It is this little green book published for the Society of Actuaries, called "Current Status of Magnetic Tape," and dated June, 1955. This contains quite a lot of information on the different kinds of tapes, recording density, organization of information laterally and longitudinally on the tape, etc., in rather useful reference form.

Dr. Grosch has just made a comment which is worth re-stating. The topic of standardization should be approached with the idea that it involves agreements to set up useful revisable specifications in a given area, usually one which has been given intensive technical development by several organizations over a period of time. No standardization program should constitute a barrier to further expansion of the boundaries of the art.

I think we might very briefly mention some of the diversities in various machines presently using magnetic tape: tape speed around 100 inches per second, sometimes 75; tape width one-half inch, five-eighths inch, three inches; information channels from 7 to 31; a wide range of track widths and recording densities; recording synchronous or asynchronous.

One important question for consideration might be the standardization of techniques for avoiding bad sections of tape, both those occurring in manufacture and those developing subsequently.

As concerns the question of making usable in one machine information which has been put on tape in form suited for another machine, there is, of course, the possibility of converting to cards and back again. This method is already at hand but is expensive and slow.

There is also the possibility of translating individual tapes from machines A, B, and C to and from a common language, but it appears that the incentive to do this will probably come only from economic pressure by users who own or contemplate owning two or more quite different types of equipment.

Finally, one can devise buffers to make tape *A* directly usable in machine *B*, etc.

I think this is perhaps all the time I have and may I again state that Datamatic Corporation is pleased to participate in this panel and will be pleased to participate and cooperate fully in any industry standardization program.

D. N. MacDonald (ElectroData Corporation): I would like to say at the outset that I hope optimism on the subject of standardization is not prerequisite for this discussion. While we are certainly in strong agreement on the efforts for standardization, it is necessary to take a realistic view of some of the problems involved. Dr. Grosch has mentioned the many variations in paper tape standards which are still with us after a number of years, and we are all aware of similar problems in punched card storage. In other fields the problem is equally serious as any of us with three-speed record players will attest. The problem of standardization on magnetic tape is considerably more severe, of course. First of all, we are dealing with a medium which is undergoing rapid technological improvements. Secondly, in distinction to some other storage mediums, the control information for its use is recorded directly on the tape and the format and type of control information is, of course, intimately tied up with the design of a particular central computer to much the same degree as a drum or core internal memory. These factors plus the demand from the field for ever increasing systems performance outweigh at the moment the demand for file standardization. However, even though we take a gloomy view of the possibility of a universal common language, there is another possibility.

Taking a parallel from similar problems in the operation of the United Nations, we can avoid the bleak prospects of a battery of mechanistic interpreters by adopting a simplified standard tape medium for communications between systems while retaining a highly integrated file peculiar to each machine. Each manufacturer would then have to provide conversion facilities only to this one *machine esperanto* instead of to all other languages and the development of local tape systems could proceed without endangering the ability to intercommunicate. While such an approach is a compromise, its adoption could readily form a basis for further standardization in the future.

M. J. Stolaroff (Ampex Corporation): Probably one of the greatest problems which plagues any new and rapidly-growing industry is the problem of standardization. As a new art is first developed and begins to unfold, standardization is out of the question, as not enough is known about the various possibilities and potentials of the art to decide upon the advantages of any given set of conditions. Furthermore, to insure that the new technology develops across the broadest possible front, and thereby makes the maximum contribution to our technology, it is important that new directions not be stifled by coercing creative talent into predetermined channels. However, the great danger arises in that, once momentum is gathered in a particular direction, and particularly if investment in tooling and production methods has been made, it becomes more and more difficult to abandon the methods already followed to adopt more standard procedures. And the longer standardization is delayed, the greater is the expense and confusion created when it ultimately becomes necessary to standardize.

Many of these problems have recently been faced in the relatively new and fast-growing magnetic recording industry. As largest suppliers of professional magnetic recording equipment, we have faced and been forced to resolve many standardization problems in the broadcast and professional sound recording field, as well as the extensive field of instrumentation recording. As a result of having large quantities of equipment in operation in a great variety of services, we are well aware of the problems that can be encountered when interchangeability of recordings is ignored. For example, you may imagine the predicament of the record manufacturer with a vault of priceless master recordings, who buys new equipment and finds the azimuth of the record and reproduce head gaps somewhat different than previously employed. Yet such a situation may well be encountered in the computer field, if many valuable program tapes are accumulated, and recording standards are changed on new equipment, or acquisition of a different suppliers' components is contemplated.

Ampex, some years ago, adopted the policv of furnishing standard track spacing for recording so that as data gathering systems grow and become more complex, interchangeability of records will always be possible. Thus once a data reducing center is established, it is possible to add more and more data gathering facilities and still make use of established analysis centers. The need for standardization in the field of flight test recording has been sufficiently recognized so that the leading suppliers of magnetic recording equipment for this field have mutually agreed upon standards for track spacing to permit interchangeability.

In the computer field, standardization appears to be even more vital, as the possibilities for applications through our industry and the need for exchange of information are potentially far greater than any other past field of magnetic recording. A firm supplying magnetic recording equipment for computer applications such as ourselves can build heads to order with special windings and special channel spacing; yet we recognize that the utilization of the equipment will be greater if desires for expansion and integrating of additional functions can be made without facing the problem of interchangeability. In particular, if the major computing systems operate with standard requirements for recordings, it becomes possible for many smaller firms to produce special equipment—data-gathering systems, data-conversion processes, and similar auxiliary devices which will allow greater flexibility and greater utilization of the central equipment. Thus by agreeing upon standards at this time, a great step can be made aiding the more rapid growth and diversification of the computer industry.

It would appear that the problem of standardization can best be approached in a series of steps, starting with fairly simple problems and proceeding to steps of increasing complexity. The first step would be to standardize on track widths, track spacing, and tape width. This would permit various tapes to be played on the same tape transport. With this much standardization, it might still be necessary to plug in different read and write amplifiers for tapes recorded in different manners, but it would still be possible to utilize the same tape handling mechanism.

The next step of standardization would involve pulse configurations, and would cover such things as packing density and pulse shapes. With these characteristics tied down, it would then be possible to use common read and write amplifiers.

The next step of standardization would begin to approach the problems of coding, computer language, etc. In this area the problem gets a great deal more complex, and may take a considerable amount of time and effort to reach agreement. However, if we start with the simpler steps first, very appreciable gains are possible, and more and more equipment will be interchangeable.

At Ampex, we have recognized the need for standards, and have formed a Standards Committee for the express purpose of foreseeing and dealing with standardization problems. The chairman of this committee is our Director of Research, Mr. Walter Selsted. Members of this committee belong to many of the national standards committees in various industries. Just as we have worked with magnetic recording equipment manufacturers in the past to achieve standards for instrumentation data and the broadcast industry, we will be most happy to collaborate with other equipment manufacturers in the computer field to resolve this problem, which if neglected, will grow ever more pressing.

J. R. Stovall, Jr. (Sperry Rand): We at Sperry Rand are very happy to cooperate in tape standardization. Unfortunately, however, we are close enough to the problem to have moments of pessimism. We agree that the motives are good, and that standardization can be achieved somehow. Beyond this general feeling, the difficult work is still to be carried out.

For purposes of clarification, let us note the two primary types of needs for standardized tapes as a communication medium.

1) There is a need for communication between computers within a company or within Government branches, often between remote points. Of course, as you might suspect, Sperry Rand prefers to solve this problem by selling only Sperry Rand computers to all points within a company or within the Government—but this solution is, I suppose, not practical.
2) There is a problem of communication between different companies or between different branches of the government. Extensive use of magnetic tape for this purpose is perhaps doubtful.

The users' reasons for desiring standardization are quite obvious. The manufacturers' reasons are somewhat less obvious. To understand the manufacturers' point of view, let us look at their exact reasons for cooperating in standardization. And, let us also look at some of the history of standardization efforts in other fields.

As for the manufacturers' reasons to cooperate, there are mainly two: Firstly, we as manufacturers are always anxious to please the customer, and the customer wants standardization. This often tends to override any other reasons we may have. Secondly, if our equipment is standardized in language, we can sell it into installations of competitive equipment, as a part of the system.

Historically, standardization among manufacturers has most easily occurred when there are a large number of manufacturers, none of whom feels that he has a possibility of dominating the market. This situation does not exist in the computer field, where there are perhaps a dozen or so main competitors, and where several are making a huge effort to achieve the largest possible share of the market. This perhaps means that an unusual amount of coercion may have to be used to produce standardization. A classic example in the history of standardization can be seen in the phonograph record field. This was standardized for years on a 78 rpm record with certain technical properties. Then after the war, the desire on the part of the manufacturers to both improve the records and to dominate the field (which motive was stronger, I will not say) resulted in our being subjected to three separate standards, $33\frac{1}{3}$, 45 and 78 rpm. Recently, in fact, 16 rpm records were introduced for the automobile phonograph market.

A final, and even more difficult problem, exists when we consider what is being asked of logical designers and engineers in standardizing on tape methods. (Again, these comments are not intended to be pessimistic, but to be realistic.) First of all, we are asking the designers of tape, tape units, and computers, to agree on some very complicated and interrelated factors in design. We are not actually asking them to standardize on what is a single and separate thing, such as a tape and tape unit. A tape unit feeding a computer is an integrated system. The reading method, the writing method, the code, the use of redundancy, the computer checking system, etc., are factors affecting the computer logic. All these and other characteristics reflect themselves throughout the system.

I think that a pertinent example to show the difficult nature of the problem is this: The particular combination of tape unit and computer chosen depends upon the particular approach of the engineers who did the logical and engineering design work. Identical results can be obtained by quite different combinations of individual type of tape, pulse codes, pulse densities, tape speed, variable or fixed item length, checking system, etc. Hence to get the engineers to agree on the *detail* system that is best, is very difficult, since many of the technical factors are obscure, and any changes agreed to will cause repercussions throughout the system with no improvement in performance assured.

Aside from these technical factors involved in standardization, the rapidly changing state of the art in this field is a considerable barrier to standardization. None of us in this meeting, I am sure, feel that a plateau has been reached in the magnetic recording art. Advances are being made at a rapid rate. So, at best, the standards will last for only a few years. Yet they must last long enough for the manufacturers to be willing to commit themselves without restricting their competitive position.

As a final comment, I would endorse two points made by previous speakers.

1) Computers have the inherent ability to translate foreign codes, and thus can provide their own compatibility. However, this is often awkward, and expensive of computer time.

2) A computer's individual code can be translated (between the tape and computer) by a *black box* to and from a separate *common language*. If this system is to be the answer adopted, then we should consider the fact that the *common language* used might well be the communications common language. We already have a joint committee organized for communication language standardization. Tape standardization effort, in this case, must be coordinated with communication language standardization.

In conclusion, I wish only to say that I hope some formal means of working towards standardization will result from this conference.

W. W. Wetzel (Minnesota Mining and Manufacturing Company): Mr. Chairman and members of the conference, I am honored to be able to represent my company at this meeting.

From the point of view of a tape manufacturer, I must agree that interchangeability of tapes from one machine to another appears to be desirable. As to standardization, it would be advantageous if agreement could be reached on the following items: tape width and the tolerance of width, tape length and its associated tolerance, maximum tape thickness, the type of magnetic coating, and the reel design.

Magnetic tape is only one small component in a complex computer. I therefore feel somewhat in a position of the manufacturer of sparkplugs who is called upon to recommend standards of automobile design to such companies as General Motors and Ford. I do believe that the responsibility for standardization is the primary responsibility of the equipment design engineer. When computing systems can be standardized I am certain that we manufacturers of components shall be happy to cooperate on tape standardization.

L. W. Armstrong (Department of the Army): However desirable redundancy may be as a checking device, it doesn't seem to have much place in the program for this morning. In the course of the remarks of the previous panelists, and particularly those of Dr. Grosch, I took my blue pencil and eliminated from my notes those things that I thought would be repetitious. Unfortunately, this has left me with nothing to say that I had planned. What I think I will do instead then is to address myself to the manufacturers of data processing equipment and to the communications industry in an attempt to drive home, from the viewpoint of one using organization, what the stakes are that we are playing for in this game.

Mr. Fitzgerald, yesterday, cited three criteria for determining whether or not there was a real need for remote data input. I believe that he implied that if any one of three situations were present, the need could be assumed to exist. These three situations were: a high time value, a central processing point, and multiple operations at a single data processing center. All three of these situations are met at most of the Department of the Army's proposed data processing installations.

Mr. Butler spoke of a requirement for about $200 million return on the present investment in data processing equipment. By a happy coincidence, $200 million is about one per cent of the total value of the inventories in the Army Supply System. This Supply System is large and complicated and it is world-wide. The system centers in some sixteen supply control points and extends throughout the Technical Service and Army structure throughout the United States, Europe, and the Far East. In providing better management of these stocks, there certainly is a high time value. The existing command structure already provides a requirement for central processing and at most of the central points there are many functions other than stock and supply control which are amenable to computer solution.

If we can improve the management of these stocks, so that their total level may be reduced by only one per cent, we come out with Mr. Butler's figure. We are, therefore, obviously playing for very high stakes. We must have and will have a compatible communications net for transmission of digital data. This is the area in which we have our greatest concern.

I will further state that in my personal opinion satisfactory devices do *not* exist today for the transmission of digital information. In closing, I will say that the Department of the Army considers this to be so important, not just in terms of the management of our resources in peacetime, but in terms of what may happen if we ever see the *blinding flash*, that if manufacturers are not willing (there is no question about their ability) to furnish such communications systems the Armed Forces will supply them for their own use.

S. N. Alexander (Bureau of Standards): I find myself in an even worse position than Mr. Armstrong. Not only have all my points already been stated in one form or another, but they have now been stated several times. Accordingly, I will use this opportunity to comment on this situation by saying that I am impressed by the surprising amount of agreement that there is at this time. This is gratifying and encouraging because it indicates that there is a reasonable possibility of accomplishing something. I believe that if this possibility is to develop into something practical and immediate, we will have to be willing to accept limited objectives for now. A push for early standardization, I am strongly convinced, is not wise. I am sure it will bring restrictions and limitations that will represent a very poor exchange for the advantages that may come from early standardization.

We need to recognize that we are just passing out of the formative period during which the present data processors were created by borrowing heavily from the techniques of the communications field. We are just beginning to design machines around components that did not exist earlier and that have been devised to serve the needs of the data processing field. I recommend that we proceed cautiously and give this new field a chance to grow along paths of its own inclination. When the natural tendencies in this field are evident from the interplay between the designer and the user, then it will be time to consider standardization. Ultimate standardization should have as a part of its basis the experience that comes from seeing some of the results of *natural selection*.

In order to provide some intermediate short-term efforts at standardization, it would seem advisable to try to have at least one input-output terminal on each data processing facility which operates in a coded notation that can be standardized and yet not restrict the characteristics of the facility to which it is attached. If this criterion can be met, I doubt if it matters greatly what form the common input-output unit takes. This follows from the fact that it will probably not endure for long because the rate of development is so rapid.

This choice ought to be geared with or related to the growing need for transmission of information between data processing facilities. If this can be accomplished, then the interim standardization would make a doubly useful contribution. Here again, it probably would not be wise to strive for unusual efficiency in providing an output to feed into a communication network as well as means to receive from a communications network. Even if we attain only a limited means of transferring information between facilities, this should be an acceptable goal for the immediate future. There is presently no assured justification for exchanges of large volumes of information on a daily basis. There are other alternatives that can be employed until we have the operating experience to indicate the volume and speeds with which information needs to flow between a related group of data processing centers. Once the medium is properly identified and specified, an appropriately engineered class of equipment will arise to meet the need. It is very important that we have operating experience on which to base our extrapolation into new equipment requirements and suggested standardization.

Thank you very much for this opportunity to express my views. Again, I must say that the surprisingly large areas of agreement encourage me to expect some substantial activity to come out of the start we have made at this Joint Computer Conference.

R. C. Matlack (Bell Telephone Laboratories): I mentioned earlier, when I was discussing the transmission problems, that a JCC-sponsored committee had been formed to explore the possibility of standardization in the data transmission field. The initial objective was rather modest and several computer people, one user, and a representative of the Bell System sat down to explore whether or not they felt something could be done in the way of standardization. This organizing committee has met and feels that something can be done in the way of standardization such as discussed by the panel speakers this morning. This organizing committee will report to the JCC that we feel a standardization committee should be formed and be composed of representatives of computer manufacturers, communications, and users to tackle the problems about which we feel there is a reasonable case for agreement. In other words, learn to walk before we run. I feel optimistic regarding future standardization in this field. Thank you.

J. Wegstein (National Bureau of Standards): Does it seem feasible to separate standardization into:

1) Standardize the form of storing data on magnetic tape for communication.
2) Standardize equipment itself.

M. J. Stolaroff: As I previously mentioned, I think there is merit in taking standardization in degrees. I believe that the items which we wish to standardize are those things which are routine, well established or understood, or relatively inconsequential. By tying down the things that are known or do not particularly matter, we then release our creative talents to tackle the more complex problems and applications. Thus standardization makes the known areas established, and we can concentrate our attention on the areas where it is more critically needed, on the growing edge of knowledge.

M. Goldstein (Remington Rand): Assuming information is represented in exactly the same manner, tape widths and reels are the same size, and so forth, how compatible are metallic and plastic magnetic tapes? (No answer.)

Conference Summary

JAY W. FORRESTER†

I WILL summarize the conference, try to estimate what the future of the computer field seems to be, and do a little lobbying for a pet subject I am trying to promote.

Measured by attendance, it was a very successful meeting. The first of these meetings, in 1951, was in Philadelphia, and 880 people attended; 1952, had 1,100 people; 1953 had 1,400 people; 1954 had 1,750; and the registration here is 1,950. Now, if you plot these, as I have done, they fall on a very nice straight line with five points, and we should therefore be justified in extrapolating this to 4,600 people in 1965—if the hotel arrangements committee wants to make its reservations.

At the first of these meetings in 1951, I also had the fun of being summary speaker and making a forecast of the future. Perhaps the only way I can justify coming back again this afternoon on the same subject is to take the position of a psychology professor of a friend of mine. When the friend went back to a college reunion, he asked the professor if he was still giving the same kind of hard questions. The professor said, "Yes, in fact they're exactly the same questions." The friend was surprised, "Well, don't all your students get a hundred by now?" The professor said, "Oh, no, they are the same questions, but the correct answers are different every time."

We have had, I think, several classes of papers at this meeting. I would like to divide them into different categories. The initial paper by Mr. Colemen I want to put in a class by itself, and I will come back to it later. We have heard five papers on the application of machines to the reduction of clerical cost, three on the subject of large files, three on remote data inputs, a few miscellaneous papers, and a continuing discussion of standardization.

On the whole, I have been impressed by the similarity between this conference and conferences held five years and also ten years ago. Ten years ago was the beginning of the design phase. Five years ago was the period when engineering applications were getting underway. Today we are talking about business applications. The one thing in common is that at none of these periods have we known what we were talking about. . . .

There is a lot of backing and filling. I don't mean this in a derogatory sense—we've been in an exploratory phase. Should we sort or shouldn't we; special purpose computers or general purpose ones; standardization, where to do it, if at all? Programming—is it a big problem or is it going to come as a package with the machines? All of these questions arise in trying to get a feel for the problem and to see what is going to unfold.

Going to the subject of standardization, I have little to add. It struck me that one speaker this morning had a very practical suggestion to standardize on communication tapes rather than on tapes for the inside of the machine itself (that is the actual working tapes). I doubt that standardization of the magnetic tapes which are intimately connected with the internal machine will progress as rapidly as the technical development in the field. In other words, any standardized tape will lag behind the state of the art by several years. In the meantime, we might standardize on something that can serve usefully even if old and obsolete. In other words, use it as we now use the five-hole paper tape even in some of the highest speed machines. The five-hole paper tape as a communication medium is very useful. We can update this one step to some newer, higher speed technique but, while doing so, recognize it as a technique that can usefully exist while technically obsolete.

We should couple communication tape standardization with the standardization of communications circuits. These two go hand in hand. I did not detect in the previous discussion, a recognition of these two being almost one and the same item. They must, I think, be talked about together because the tape standardization we get may simply be a way of recording the standard signals that are going on the wire networks that tie machines together. If we are able to communicate into a standard set of communication networks, then we've gone a long way towards the kind of standardization I think most of the people in the field really need.

We have had several discussions of the large file for business purposes. I can't add to this except to give my own opinion—the problem is going to be solved, solved fairly successfully, and in two to three years, at least on a development basis. The people who are taking the long range view of their planning might be well advised to think in terms of the kind of systems that could be set up around a reasonably successful file for large amounts of business data.

Now, while we are predicting, look at the rate electronic developments are progressing. We have had a technical advancement of about a factor of ten per year in this field for the last decade. Speed, storage capacity, reliability—one of these has gone up by a factor of ten almost every year, and is the reason why the field confuses people who now enter it for business applications. This rapid evolution is still in progress, and it appears that presently foreseeable techniques can insure a factor of ten improvement per year for a few more years.

† Division Head, Lincoln Laboratory, and Director, Digital Computer Laboratory, M.I.T., Cambridge, Mass.

The users, the people in business, plead for stability. But I don't think they really want stability. The machines are not yet good enough for them; and, their own ideas, in fact, aren't stable. They don't known precisely what they want. I think this is an evolving field, and there is no more stability in the business applications end of it, at the moment, than there is in the electronic end.

It was only ten years ago that computer speeds were the speeds of relay devices, and they were measured in tenths or hundredths of a second. A little later we had gas tubes with speeds in milliseconds, and then vacuum tubes working in microseconds. Now we have people in solid-state physics with transistors and magnetic materials who are talking about measuring time in millimicroseconds. It is obvious that we need a new measure of time. We've just about worn out the second, and perhaps we next should start using the jiffy. The jiffy you know has been defined by some physicists as the period of time it takes light to travel one centimeter.

As an industry, the field is growing very rapidly. Several hundred million dollars worth of electronic computing equipment has now been sold. Even more impressive is the fact that the backlog, the unfilled orders for this equipment, equals or exceeds the total amount delivered thus far. This, along with other observations, leads to the conclusion that the computing capacity in this country is doubling about every two years. I think there is every reason to believe that it will continue to double every two years until we reach several hundred times the present computing capacity.

We can divide electronic computer evolution into five year periods, each period with its paramount significance. 1945 to 1950 was the period of electronic design. From 1950 to 1955, attention has been focused on the solution of scientific and engineering problems. 1955 to 1960 will encompass the upswing in the commercial data-processing applications—the things that we have started to talk about at this convention. 1960 to 1965 will probably mark the shift of major attention to the use of digital computers as the central elements in real-time control systems.

In anticipation of the 1960–1965 period, computers already are calculating punched-paper and magnetic tapes which are fed into newly designed machine tools for high speed, accurate automatic control of the 3 to 5 control axes. In recent news releases from the Air Force you have seen some of the story on SAGE, the Semi-Automatic Ground Environment System. General purpose digital computers, as outlined in these releases, are to be the nerve centers for tieing together the flow of information in our forthcoming new air defense system. This type of control system, we can assume, will develop further into a high speed automatic control and regulation of future civilian air traffic. Better air traffic control will come none too soon. The papers this morning say that there are now firm orders from the commercial airlines for 100 jet transport planes—the kind that will fly New York to Los Angeles in four to five hours. Delivery will begin in 1959.

Further on real-time control, consider the chemical plants and oil refineries. The trend has been towards ever increasing use of servo-mechanisms and automatic controls. In the last thirty years the automatic controls in an oil refinery have risen from a negligible part of the capital cost to some 15 per cent of the investment in a refinery. The kind of refinery I'm talking about costs $100,000,000, and includes $15,000,000 worth of automatic controls. I believe we will see digital computers as controllers and monitors of operation in these plants to permit closer control, higher speed chemical reactions, larger outputs, and a better product. In many of these plants one would need to increase the value of the output by only a fraction of one per cent to amortize one of today's million dollar computers.

Another real-time control application of computers which is being explored is in the management of electrical power generation and distribution. In the past, computational studies of power networks have been limited to the use of the AC calculating board for the solution of losses and stability of the power line network. As systems grow and power generation sources diversify, the problems become even more complex. In a combined hydrothermal power system such as found in the Pacific Northwest, the economical generation of power depends on the locations of the generators, the destination of the power, the cost of fuel which might be used in the thermal plants, the efficiency of each plant, the availability of reservoir water for hydrogeneration, the coupling between hydro plants (because the water that flows through one becomes the supply of water to another further down stream), the snow depths in the mountains and therefore the probable summer water supply, and the necessity of maintaining river navigation depths at all times—(you must not shut off the water and collect it all behind a dam or the salmon can't swim up stream, and the ships can't move). Such power network analysis is thereby expanded to the nonlinear equations describing the economic and statistical aspects of day-to-day electrical load dispatching. The avoidable electrical transmission losses alone in such a system are great enough to pay for a computer installation when people learn how to use it. Already, the data sources for that kind of system mostly exist. Telemetering over the power lines themselves to obtain information and to exercise control is a fairly common practice.

Now that we have a decade of past in this field, as well as a present and future, we have a full cast of characters on the stage of computer development. We have those who try to picture the future; we have those struggling to make the present successfully approximate past promises; and we have those who are trying to keep the whole affair financially solvent. It is a little like the relationship between the neurotic, the psychotic, and the psychiatrist. The neurotic dreams of castles in the air; the psychotic lives in them; and the psychiatrist collects the rent.

Now, let's return to the paper by Mr. Coleman at the opening session. I have heard no discussion of it among the people I've talked to, and yet, to me, it was by far the most important and challenging talk that was presented. I don't have direct quotes; but from my notes, he said, "as a data processing tool, the computer is just another in a long sequence of techniques to cut costs." And most of the papers here were devoted to ways of cutting costs, and often displayed uncertainty of success. Mr. Coleman further said, "as a tool for helping to reach higher-order management decisions, the computer shows the brightest hope for the future."

And herein lies, I think, a distinguishably different and important field of application—the use of computers by management for processing data in ways not now attempted by human clerical systems. In spite of the confusion, the difficulties, and the problems that do exist at the present time, the routine conversion of clerical work to machine execution is relatively straightforward—at least where it is economically feasible. Mechanizing present clerical procedures is not, however, the most exciting or most rewarding of the applications to which machines will be put in business. Looking a step beyond the present, one can see that the new electronic techniques will lead to changes in our management structures and functions. Many of the routine decisions that are now thought to require human judgment will be recognized as amenable to automatic execution. Better information and new forms of presentation will then help management to make the decisions truly requiring human judgment. Perhaps we can't define what we mean by "truly requiring human judgment," but we can keep pushing the threshold back and using management talent for the higher and higher level creative contributions which we haven't yet learned to describe and to mechanize. A great deal of present industrial management time is expended on tasks that could well be mechanized. Leadership is thereby lost from more productive purposes. Just as automation places a new premium on workers' skill at the factory production level, so, I think, will computers place a premium on management skill at the planning level. Given new tools, management will be obligated to use them. Moves such as the guaranteed annual wage to stabilize the industrial cycle will force management to use every facility at its command to smooth and predict these production demands.

Automatic production methods have reduced the cost of the goods which we, as consumers, buy. Over the last hundred years the productivity in this country has gone up by a factor of six. In other words, one sixth as many man hours need be expended for an average essential of life today as a hundred years ago. This is real productivity, independent of the money level and inflation. I expect the next big point of attack in industrial productivity will be on the problems and costs of distribution. About 50 per cent of the retail price of many manufactured items is directly traceable to the cost of distributing the product after it leaves the factory.

As an example of a future trend, there already are some who believe a reorganization of distribution methods and the use of electronic computing equipment for the control of inventories, distribution, and factory production scheduling could reduce the retail price of items like refrigerators and electric stoves by as much as $200. You won't save that much by cutting the existing clerical cost in a business.

The business activity of the country at the present time is at the level of about 360 billion dollars of gross national product. Many persons expect this to expand by 1965 to 500 billion dollars a year. Such an expansion would be a continuation of the trend which has brought the United States (which has only six per cent of the world's population and seven per cent of the world's land) to a level of 50 per cent of the world's factory output. In doing this the industrial capital investment per worker has increased from $500 a century ago to about $15,000 per worker at the present time.

Simultaneously, an important change has occurred in the family income distribution pattern for the country. Traditionally, in the past, this has been a triangular distribution with a small number of people receiving the highest incomes, and progressively more people in the income brackets below, with the largest number of people in the lowest bracket. This has been the past pattern for the world. This distribution is changing in the United States to a diamond shaped distribution with a few people at the top and the bottom and the largest number of families in the middle income group. This new income pattern means a larger excess of income over the cost of the essential necessities of life. This larger "disposable income" (as the economists call it) will lead to a greater competitive effort to attract these discretionary funds. It will lead to greater efforts in market research and still greater striving for higher productivity of goods per man hour of labor expended. The trend toward scientific logical industrial management is increasingly evident. The organization of the Operations Research Society and the Institute for Management Sciences in the last few years is indicative of this trend.

As we build this kind of a world then let us be sure that we will be able to look back on it with pride and satisfaction. Let us take care not to be like an associate of mine who arrived in a small airport served by only one airplane a day. He arrived in time according to his ticket; but the ticket was made out wrong. He was actually four or five minutes late, and the plane was just leaving the runway. The airline was quite distressed over the error in the ticket, and they sent up a radio message to have the plane come back. It landed. He was feeling very important to receive this special attention and was very self-satisfied as he walked up the aisle until he heard a lady behind him say, "It wasn't worth it."

Or, more seriously, let me read you some quotations[1] from the acceptance speech by W. L. Everitt of the University of Illinois on his receiving the 1954 Medal of Honor at the annual IRE banquet. The title of his talk was, "Men Who Play God." Dr. Everitt said, in part: " . . . But in a different sense, do you realize that you, too, the members of the Institute of Radio Engineers are 'men who play God.' . . . Other engineers have created machines from the materials of the earth, which supplement man's muscles. But the electronic engineer has given and is giving to these machines sense organs, nerve systems, and reasoning power. . . . Every day we hear of more and more machines which can supplement or replace man's brain power and sense organs as well as his muscles. As someone has said, 'Machines are becoming so human they can act without using any intelligence. . . .

". . . A wise man said, 'God cannot be solemn or he would not have blessed man with the gift of laughter.' We have not yet built into our machines a sense of humor, although some of them do make us laugh. . . .

". . . This world of automation we are building is our world. Be sure that when you play God you do it reverently, thoughtfully, and honestly with full attention to the fact that you are building not only machines but a world in which you and your children are going to live."

All this leads us to some of the social and public relations' responsibilities of our new and growing industry. Look around you. We recently had a series of Congressional hearings on automation by the Subcommittee on Economic Stabilization of the Congressional Joint Committee on the Economic Report. This developed from Congressional concern over the social implications of new materials and data-handling devices. Furthermore, there is a very real fear of science and technology in the minds of large segments of our population. I heard a psychologist the other day discussing some word association studies made with elementary school students. He was very distressed by the results. He found that the public, as represented by school children, associate science and scientists with such words as evil, villain, and atom bombs. These are the things that come to mind rather than the associations you would prefer.

There is need for a better public understanding of computers for many reasons. The industry itself needs more designers. It is going to need people who are able to use, to operate, and to maintain these machines. In the papers that we have heard on banking and credit card accounting, it is clear that the public needs to understand not only how some of these new methods work, why they should be used, and how they will give a better product, but people need also to understand how they must have a certain amount of forebearance and permit some changes from our present systems of accounting, billing, and sales recording.

I propose, then, that the only place that we will be able to find all of these people (the future engineers, the operators and other nonprofessional workers, and the public) is to go into the high schools of the country. Graduate school is much too late a time to introduce the universal field of computation into our academic institutions. It is important that undergraduates, probably in the freshman year, be introduced to the logical preciseness of digital computer programing. As an exercise in preparing complete unambiguous descriptions, programing has a universal value and is well within the grasp of the undergraduate or even the high school student. It is a misconception to believe that digital computers can be understood only by the engineers and mathematicians. I recently had an opportunity to observe an effort to find, in the Boston area, forty persons to be trained as computer programers. We gave special attention to the basic traits required rather than to the particular branch of formal education. The work requires intelligence and logical reasoning; and for what we were doing, it seemed advisable that recruits should possess aptitude for spatial perception. Accountants, actuaries, and business administration students qualified. You might be interested in this example. One of the highest scores on the apititude tests was made by a music major from a nearby women's college. The young lady came to work. The initial reaction in an engineering organization was one of raised eyebrows when she came around to her supervisor and asked what the words "sine" and "cos" meant. Within six weeks, however, she was making improvements in computer programs that had been worked out previously by experts. Computer programing requires interest, enthusiasm and a certain set of basic traits which don't necessarily go with any particular line of formal education.

In the next two decades, automation in the factory and electronic information processing in the office will free many men and women from their present types of work, and it will be necessary to attract many of these to the design, construction, maintenance, and management of electronic information-handling systems. We must reach into the high schools and to the public with the message that these new developments are understaffed and crying for more people rather than their being the bugaboo that will create unemployment.

Gabriel Hauge, who is economic advisor to President Eisenhower, three weeks ago in this room gave a talk at

[1] Proc. IRE, vol. 43, p. 3; January, 1955.

the Boston Conference on Distribution in which he outlined the President's economic philosophy. One of the cornerstones of this philosophy is the anticipation of an expanding economy and a future shortage, rather than surplus, of labor. It is important, therefore, that devices continue to be developed to improve our productivity per man hour. If new equipment is to be fully effective, it will require better public understanding.

Now, if you don't believe that computers are of interest to the high school level, look about you. I will cite two instances that have come up recently. We have been approached by students who want to make computer exhibits for the science fairs that are held throughout the country. As you know, students build science exhibits and have competitions, and the winner of the school goes to a state fair, etc. They want computer material in the science fairs and for school assignments. I will read you a letter which came recently:

"October 3, 1955

Dean of Electronics
Massachusetts Institute of Technology

Dear Sir:

I am in the eighth grade, and I am doing a report on "Man's Thinking Machines." I would appreciate any pamphlets, or any other information that you have on the electronic brain, or any like instruments that are thinking machines.

Thank you very much. You can send the information to: Jon Leon, 211 Pine Point Drive, Highland Park, Illinois.

One of my future ambitions is to go to M.I.T. because I want to be an engineer.

Yours truly,
Jon Leon"

We had a real problem. I am at a loss to know how to answer that letter adequately. We did the best we could. We sent him information, but I don't think it was a satisfactory answer. There isn't, so far as I know, adequate material, suitably written, to answer that kind of letter. In this, the data-processing industry is lagging behind other industries. Other people are doing more.

Let me read to you from an editorial on the back of a recent investment news letter.[2] It's a little long, but I'd like to read it all:

"Time and again we have been reminded by educators and industrial leaders that we are not training enough new *scientists* and *engineers* to keep pace with world scientific development—But what to do about it?

Well, a significant step in the right direction has been taken by a group out in Chicago known as *"Industrial America, Inc."* They are sponsoring the distribution of *American Industry Educational Hobby Kits* which are produced by such well-known industrial concerns as *Radio Corp. of America, Taylor Instrument Co., American Optical Co., Bauer and Black, Gemological Institute,* etc.

The world famous *Museum of Science and Industry* in Chicago has collaborated on the design and creative side of the project and each kit contains an illustrated booklet prepared under the guidance of *Encyclopedia Britannica.*

Actually, these kits are scientifically planned *do-it-yourself educational toys* designed for the 8 to 12+ year age bracket. They pack a terrific wallop of intriguing interest to the youthful mind, and enable children to really delve into the major sciences on their own.

The *electronics kit,* for example, contains all the necessary parts for youngsters to learn how to transmit and receive messages on their own home-built radio sets.

With the *weather kit,* boys and girls can build all the instruments needed for their own complete weather station and learn how to observe and forecast weather.

The *medical kit* contains X-Ray Eddie, an inflatable plastic anatomical model that enables children to see how the body's blood stream, breathing, digestion, and bone structure functions. Youngsters can even take each other's blood pressure with instruments made from the kit.

Priced reasonably and available in many toy and department stores, these *"American Industry"* kits should be very popular items during the coming Christmas season. Also—the scientific interest they generate may well fill our Engineering Colleges in the 1960's."

To follow up that editorial, I went yesterday to F. A. O. Schwartz, the toy store which is three or four blocks from here, to look at such hobby kits. They are really quite impressive! They're well done! I would judge from the prices, that run around $20, that they are completely on their own financially. I think the price is a matter of importance. They are not inexpensive toys, and they are not million dollar computers. If you look at what they do in their particular fields, you find we could do the equivalent.

Since I have not run into any opposition to the importance of contacting the high school level. I expect that most of you would agree that something should be done. The problem is, what will we do to keep from dropping the subject with everybody agreeing but nobody doing anything. I would therefore like to make a proposal that we arrange to have a team of full-time men—perhaps five or six, one donated by each of the several large industrial concerns in this field—who would work together, not as ordinary professional committee members meeting occasionally but as full-time working men with no other goals for a period of a year, and that they do this under the auspices of the Joint Computer Committee that sponsored this conference. I have not cleared this with the Committee, but I've talked to some of the members who seem to think it is a good idea. Get some people together who are interested in this educational program and give it to them as their full-time job, give them sponsorship and support from industry, and let them go to work.

The toy kits, I have mentioned (I think it is only necessary to provide material, designs and ideas and that it would be best to let somebody already in the field, like Industrial America, Inc., carry out the manufacturing and distribution). Another approach is to make arrangements with a good author who knows the high school audience level and help him to write on computers, automation, the kind of work the people will do with these machines, and the impact they will have on society—in other words, answer the questions that the youngsters are asking. There are some very good books in the science kits along this line. Also the group could encourage exhibits for our science mu-

[2] Paul Talbot, "The Back Yard," United Business Service, Boston, Mass.; October 31, 1955.

seums. I am very sure that the science museums would enthusiastically participate if they had a little encouragement and some sources of information. Particularly, the group should plan to help the high school science teachers in various ways (narrated film strips, experimental supplies for lab work, etc.).

I think one of the most important educational contributions would be do-it-yourself instructions that would allow the youngsters to make things representative of computers and computing circuits, and automation and controls, out of the parts that are readily available to them without expenditure for hardware. For example, supply instructions for electronic counters made of discarded radio parts. Or, the country is strewn with perfectly good relays out of automobile voltage regulators that have been discarded because the servicemen don't know how to test them. I watched a serviceman work on my car. He tried to adjust a 12-volt system to $7\frac{1}{2}$ volts by reading the 10 volt scale of a meter that he had connected to the 50 volt range—a fine set of relays to build toy computers from when he finished.

Now, you ask how to finance such an educational program. There are many ways, but let's take the straightforward approach of considering it a normal and important part of doing business in a new field. There are about 30,000 high schools, and they have in them about 7,000,000 students. I went through a list of industrial companies last night. By picking about five companies which are predominately in the electronic data-handling field, one totals a combined yearly sales of 1.2 billion dollars. I would suggest that these companies might well afford to take one-tenth of one per cent (0.1 per cent) of sales for a joint industry educational program at the high school level. This makes a yearly budget of $1,200,000; it is $40 apiece in each high school. Compare this with the two, three, five per cent, or higher, that each of these companies is spending on physical research to develop equipment to be sold to the bewildered public.

C. B. Caldwell, Vice-President of Sears, Roebuck for Personnel and Employee Relations, pointed out recently that we spend millions to eliminate friction in machines, but practically nothing to eliminate friction between people. We are talking here about eliminating friction and misunderstanding.

Touring lecturers now supplied by a few companies are not enough. They don't reach enough people nor the small places. They don't reach the rural communities. And let me point out that the rural communities are important. A recent analysis of names in "Who's Who in Engineering" was made on the basis of the percentage of each college's engineering graduates included. Colleges in the rural areas rank high. This survey showed Dartmouth at the top, second, University of Nebraska; fourth, Kansas; fifth, Missouri; ninth, South Dakota State; tenth, Missouri Mines; fifteenth, Iowa; and eighteenth—MIT. I think this is significant. The people that are on the country's farms are important to us in new technical fields.

You can look at the relationship of the computer field to the farm in an entirely different way. Perhaps we can help solve another problem that has been brought on by technological changes—the farm surplus problem whose solution has been sought by legislation rather than education. Fifty per cent of our farms produce ninety per cent of our food. That means that fifty per cent of our farms produce ten per cent of our food; it's the fifty per cent that produce the small end that make the farm problem. The others can make a financial success of farming without controls and price supports.

Look at the relationship between the computer industry and the farm support program on a dollar basis. The computer industry—again the four, five or six companies which are major producers of computing equipment—pay about $90,000,000 a year in taxes. In federal expenditures, about one per cent of the federal budget goes to the farm price support program. That means that the companies in this industry are putting one per cent of $90,000,000 or $900,000 a year into the farm price support program. That is $30 per high school.

$30 to $50 per high school per year should carry a very successful program. You see, the kind of money it takes is relatively small compared with what we spend in research, compared to supporting farm programs that exist because of lack of education to lead people into other activities, compared to other public relations expenses, and indeed compared to advertising costs that try to attract engineers from the end of an educational pipeline which we are making insufficient effort to fill at its source.