The activities of the Division of Accounting Operations of the Bureau of Old-Age and Survivors Insurance have been referred to, upon occasion, as one of the biggest bookkeeping jobs in the world. Whether this condition is true or not depends on how it is measured. Certainly, by the yardstick of costs, it would not appear that the job is at all near the biggest. On the basis of comparisons similar to those made in business, the gross costs of operations represent only 3/4 of 1 percent of total gross income. Furthermore, the system shows no signs of becoming the biggest in terms of cost. In a ten year period, an approximately 50 percent increase in basic work loads has been absorbed with no significant or compensatory increase in personnel. It is felt, therefore, that if the job qualifies as being one of the biggest in volume, it also might qualify as one of the smallest in terms of proportionate costs. However, regardless of how it is measured, the job represents a real challenge to those who administer it and to those who would furnish it with suitable electronic paraphernalia. For these reasons, the following problems are presented with pleasure at the opportunity and with confidence in the know-how represented at this conference to solder together whatever combinations of wires, tubes, diodes and transistors that are found to be necessary in each case.

The largest area of mechanical operation within the whole accounting system is in the processing each quarter of approximately 60 million entries received in random order to approximately 107 million accounts which are established in numerical order for ready reference. The end result of this work must permit rapid access to any one of the approximately 107 million accounts at any time except the instant at which a particular account is being posted. Such access is necessary in completing the approximately 755 thousand references made each quarter for such purposes as: processing claims for benefits, issuing statements of account, and making adjustments.

The ideal situation, of course, would be one in which each of the 60 million incoming items each quarter would be incorporated with the related account at the instant the new item is presented to the system. Key depressions by which punched cards now are produced instead would cause pulses to flow over wires to a device which would summon forth the designated account. This device would compare and post the entry or separately record it for clerical investigation if absolute identification of the entry with the account were not possible. Although, we often have seen yesterday's fantasy become today's reality when electronic principles have been established and applied, we cannot wait for as long as might be required to produce equipment suitable to this ideal. Hence, we must contain ourselves with projects of less ambition while the science progresses to this ultimate destination.
A sorting problem of considerable magnitude, therefore, is involved in arraying 60 million items each quarter for comparison with and entry into 107 million established accounts. Each of the 60 million items contains a nine digit account number and 61 digits of satellite information. Six words probably would be necessary, therefore, to express each of the 60 million items in electronic notation. We have not found, as yet, any way in which this sorting work might be done with any of the existing general purpose computers at a cost comparable to that of the rental of punched-card sorters. One of the presently most capable general purpose computers is reported to sort at speeds 5 to 10 times faster than that of a punched-card sorter. However, the cost factors greatly exceed these ratios. Attempts have been made to see how the cost of sorting by a general purpose computer might be reduced by some combination of other operations with the sorting. Unfortunately, it has been found that such possibilities so far have not been sufficient to surmount the cost disadvantages involved. It appears, therefore, that before sorting can be considered to be an efficient electronic process, either of two developments must occur. It must become possible to feed and extract data from general purpose computers at rates many times faster than at present or to have a low cost special-purpose computer specifically designed for sorting.

Although such sorting facilities are highly desirable, the present lack of them might not preclude some application of electronic equipment to this problem of entering 60 million items each quarter into 107 million accounts. A system which combines the advantages of punched cards and magnetic tapes might be found to be the most suitable answer for awhile. Under this arrangement, the data would be introduced and retained in punched card form until it had been sorted by account number and made ready for the actual posting to the established accounts. At this point, converters might be used to transfer the data from punched cards to magnetic tapes. Preliminary studies have been made of this possibility. Conclusions at this point are by no means final, and additional study is necessary based on more recently available operational data on card-to-tape converters and general purpose computers. However, some observations might be stated with reasonable validity at this time.

Card-to-tape conversion is an economical procedure if the operations performed thereafter in an electronic medium accomplish results at significantly lower cost than the punched card medium permits. The foremost requirement is that the cost of the card-to-tape conversion be absorbed. This requirement involves consideration of the cost, number, extent, and variety of the clerical and machine tasks which might be performed in one pass through an electronic computer but which would require separate and extensive treatment with punched card equipment, or by clerical workers. An additional requirement, insofar as the data processing problem previously discussed is concerned, is that the means be provided for frequent and random interrogation of the tapes.

Our preliminary studies show a possibility that two large punched card machine operations might be wholly combined and two clerical operations partially combined into one electronic computer operation after a card-to-tape conversion. The mechanical operations involve the collation and posting of the 60 million items to the 107 million accounts. The clerical operations involve the reconciliation of minor discrepancies in the spelling of names and
certain transpositions in the account numbers, and the notation and reconciliation of cases wherein females have changed their surnames by marriage or divorce after their accounts were established. Whether or not it is desirable to convert from cards to tape to accomplish these combinations of operations will depend on the results of further cost studies and on the finding of suitable means of obtaining high speed random access to data stored on electronic media.

The reason a means of high speed random access is so important to a substantial conversion to the electronic medium is that approximately 755 thousand random references are necessary each quarter to the data contained in the 107 million established accounts. Approximately 400 thousand of these references are made in connection with claims for benefits and 230 thousand for the issuance of statements to persons who request information concerning the amounts recorded in their accounts. It is necessary that this type of work proceed rapidly in order that the public might be served properly. Consequently, work cannot be scheduled and arrayed for reference to the file of accounts in blocks large enough to provide a considerable density of reference. Contacts with the file at widely scattered intervals are necessary. The time limits on the accumulation of work for reference purposes cannot be extended for much more than one hour per block. In addition to this situation, the references which are made involve the removal from file of a set of media, at present in the form of punched cards, for the mechanical preparation of abstracts of accounts and benefit computations for claims, and for the mechanical preparation of statements of account. The file to which these references are necessary contains, at present, approximately 20 billion alphabetical and numerical characters of information in punched card form. It appears that the present inability to obtain the discrete use of data when it is stored on reels of tape would seriously handicap the type of reference work just described. It would be necessary to move many thousands of reels in and out of the tape file daily. In numerous cases, the same reel might be required upon more than one occasion during the same day.

With a large number of inputs to a computer, it might be possible to work under such conditions provided the time required to load and unload the tape feeding devices could be substantially reduced. One company reports that a skilled operator can load or unload its machine in one minute, and that it is attempting to reduce this time to ten seconds. The success of this effort might make it possible to conduct these reference operations by manually carrying reels to and from a computer instead of relatively small quantities of punched cards to a printing tabulator as at present. However, it also would be necessary to have quite durable feeding mechanisms to withstand such constant use. There also would be a need for definite assurance that the frequent daily handling of the reels would not create serious problems of tears or kinks, or of dust and dirt becoming imbedded in the tape. For these reasons, the prospects of being able to conduct large scale random reference operations without equipment specifically designed for that purpose are not viewed with substantial optimism at this time.

A high-speed random access external memory of capacious design also might provide the means of applying electronic data processing equipment to another important area of reference work. The Division of Accounting
Operations maintains a file containing, at present, approximately 138 million flexoline strips showing the identities of holders of account number cards. Each strip contains a 9 digit account number, 25 digits for the name, a 3 digit Russell soundex code number, and 6 digits for the date of birth. Approximately 36 thousand references are made daily to this file of flexoline strips, principally to identify applicants for duplicate account number cards and wage earners who were reported under incorrect account numbers. This type of file was adopted and has been retained because it offered, and still offers to the present time, the highest speed of access to such information. For example, over 15,000 of the daily references are made to determine the account numbers of individuals so that duplicate account number cards can be issued. If the applicant for the duplicate card has given the same identifying information as he furnished originally, it usually is possible to locate the correct number in a matter of seconds.

Clerks who make the references to this file are required to do considerable walking, standing, reaching, and bending, notwithstanding the fact that the work is blocked to obtain the maximum density of reference. The conversion of the data in this file to a high-speed random access external memory appears attractive, not only from the standpoint of economy, but as a prospect for the elimination of a laborious type of work. For these reasons, we should like to see more attention given to this type of development. At the moment, we know of but two efforts which have been made in this direction. The one is the Notched-Disk Memory invented by Dr. Jacob Babinow of the National Bureau of Standards; the other is the R.A.M. of the Potter Instrument Company. Both of these devices now are being studied to determine their possibilities. If other equipment of this nature exists, or is being planned, we should like to have whatever information is available.

Two areas have been found in which applications of electronic data processing equipment do not depend on the availability of equipment for high-speed random access to data stored externally. The one involves the calculation of the primary insurance amount of each claim that is processed; the other involves the development of the statistics which are necessary in certain operational, program planning, and other activities of the Bureau of Old-Age and Survivors Insurance. In the processing of claims, the computation of the benefit payments are based upon the claimants' work histories. The Social Security Act, as amended, requires the consideration of different base periods, the inclusion of credit for military service, and the use of a number of different formulae in the computation process. There are ten possible methods of computation; the one which gives the highest benefit must be selected. In this work, a computation card is punched for each of the pertinent possible methods of computation for each case. These cards are processed in an IBM Type 604 Electronic Calculating Punch Machine at the rate of 100 per minute.

A considerable amount of thought and attention has been given to the possibility of using electronic data processing equipment in the statistical operations of the Bureau. These statistics are used in analyses of the old-age and survivors insurance program and in administrative planning, and are used also by certain other governmental and private organizations for
general economic and demographic studies. The statistical operations involve the quarterly and annual processing of millions of punched cards obtained as a by-product of the regular accounting operations. In spite of the fact that a great deal of the basic source material is obtained as a by-product of the accounting operations, a substantial expenditure both in manpower and standard punched card machine is required to convert the basic cards received to final statistical data. Because the Bureau has not been able to obtain all the potentially available and required data, due to limitations of present equipment and time schedules, the Bureau is investigating the possibility of using electronic equipment to overcome these problems in the statistical program.

The investigation was started in cooperation with the National Bureau of Standards as part of a survey which also covered the investigation of possibilities in the accounting field. A number of test runs indicated some possibilities of money and time savings in the statistical operations of the Bureau of Old-Age and Survivors Insurance. However, these tests were run with but a few thousand accounts. Consequently, the results were not deemed conclusive enough to estimate savings which might be realized under actual conditions, when the number of accounts to be processed is substantially greater than in the test runs. Hence, a decision was made to conduct additional studies involving several hundred thousand punched cards. The Bureau of Old-Age and Survivors Insurance is arranging to have the Bureau of the Census perform two of the former's largest statistical tabulations on the Census UNIVAC machine. This machine will not be available for these tests until the Spring of 1953. In the meantime, staff members of the Bureau of Old-Age and Survivors Insurance are programming their statistical operations and will attempt to have the programs tested, so that the Bureau will be ready to proceed with the actual statistical operations as soon as the computer time becomes available.

In one of the two tests mentioned, approximately 300 thousand punched cards would have to be processed by conventional methods to obtain information on the economic and personal characteristics of about 100 thousand individuals with social security account numbers. In this operation, data are summarized for each individual. In addition, approximately fifty statistical tables are produced. This particular operation is similar to another annual statistical tabulation of the Bureau which involves ten times the number of cards and individuals and over 100 statistical tables. This latter operation comprises more than half of the total expenditure by the Bureau for compiling statistics. The Bureau is going ahead with the smaller statistical operations test on the assumption that the results can be used in making comparable estimates for the larger statistical operation.

In the second statistical test, approximately 25 thousand punched cards containing information on selected employers reporting under the old-age and survivors insurance program would be processed to produce approximately ten statistical tables by conventional methods. This operation is performed annually in this Bureau for approximately 3 million punched cards and is financed jointly by the Census Bureau and this Bureau. The results are a publication "County Business Patterns," which you might have
seen. Again, it would be necessary to draw inferences about the timing and cost for completing the latter operation on the basis of results obtained in the test.

This Bureau also has found that high speed printing equipment would be necessary to any large scale conversion to electronic data processing equipment. However, it appears that such requirements have become rather widely recognized and need not be argued for in this paper. In view of the comprehensive treatment given to this topic at the computer conference in New York this last December, it appears that satisfaction of this need lies at hand or at least is within reach in some form which might not require large reconciliations of expectations with the facilities which are offered.

Electronic scanning devices which could read pages of information and activate mechanisms to form characters and numbers in the form of holes in punched cards or magnetic patterns on electronic media also could be used somewhat extensively in the operations of the Division of Accounting Operations. A considerable number of the 60 million items received for credit to the accounts each quarter originate from returns of employers who use punch card equipment, typewriters, or addressograph machines. The printing on many of these returns is quite legible throughout. In addition, the information is spaced and aligned in a manner which probably would not require shifting beyond the range of the scanner. In view of the large turnover of personnel in key punch operations and the expense of key punch operations both in punched card and electronic data processing, a development of this kind would be most favorably received. For this reason, we are watching with considerable interest and anticipation a few developments which appear to have substantial prospects of application in this field.

In conclusion, it is felt that the requirements which have been stated in this paper are not altogether unique. Comprehensive market surveys by interested manufacturers probably would uncover quite similar needs in numerous other governmental agencies and in commerce and industry. For these reasons, this paper is presented with the hope that it might prompt such inquiry and possibly cause other potential users to come forward with descriptions of their requirements. Such information, when gathered, might form a basis for coordinated effort leading to the earlier satisfaction of a wide area of common needs.

Acknowledgment:

Acknowledgment and appreciation is due to Mr. S. D. Hearn, Division of Program Analysis, Bureau of Old-Age and Survivors Insurance, for his having prepared the discussion of the statistical problem included in this paper.

References:


THE PROCESSING OF INFORMATION-CONTAINING DOCUMENTS

George W. Brown and Louis N. Ridenour
International Telemeter Corporation
Los Angeles, California

The Problem

General Remarks

Much recent attention has been paid to the promise that digital computer techniques will simplify, speed up, and cheapen various sorts of large-scale information processing in business and industry. Only a few actual applications of this sort have so far been attempted, probably for two main reasons. First, punched-card machines are already so highly developed that, despite the potentially greater speed of the newer electronic techniques, it is difficult to introduce novel ways of accomplishing what can now be done by standard business machines. Second, and probably more important, appropriate input and output equipment to couple the world of the digital computer to the world of men often does not exist. To use a computer for scientific or engineering calculations, it is sufficient to provide it with input in the form of a device for reading punched cards or tape, and an output in the form of an electric typewriter or a card punch. Most existing computers have terminal equipment no more sophisticated than this. To use such a machine for accounting purposes, however, requires a far more imaginative solution of the input-output problem; a satisfactory solution can be achieved only on the basis of a deep understanding of the nature of the accounting activity that is being mechanized.

Thus, before the automatic processing of information can be fully successful, means must be provided for imparting the pertinent information to the machine in a form suitable for automatic handling, and for extracting the processed information in useful form. In many practical situations, the problem is made still more difficult by the fact that the original documents containing the information to be processed themselves have a significance, and must be physically handled, routed, or sorted on the basis of the information that they contain. It is an application of this sort that we wish to discuss.

While documents are often used only as vehicles for information, and can be replaced by other documents carrying the same information, there are many other instances in which the original document itself possesses a logical or legal significance, so that it must be preserved throughout an operation which may be rather complicated. In the case of a bank check, for example, the original bit of paper constitutes at all times the legal evidence of an obligation which is fully discharged only when the check is cancelled and returned to the man who drew it. Bank checks must be physically conveyed from hand to hand until they are deposited in a bank; thereafter they must be taken through a series of sortings and concurrent proof and bookkeeping operations, until the bank on which each check is drawn returns each check to its originator.