COMMERCIAL APPLICATIONS — THE IMPLICATION OF CENSUS EXPERIENCE

James L. McPherson
Bureau of the Census

The Bureau of the Census began operating a Univac System in April 1951. We temporarily stopped operation at the end of December 1952. During this period our Univac was housed in the factory in Philadelphia where it was built by the Eckert-Mauchly Division of Remington Rand Inc. The first of this year (1953) our Univac was shipped to Washington. It is now being re-assembled and we hope to put it back in operation within the next two or three months.

As soon after our Univac was delivered as operating personnel was partially trained which was about June or July of 1951, we introduced a seven day per week, twenty-four hours per day schedule. This schedule was continued throughout our operation of Univac.

During this period the Univac performed a variety of what we call "small jobs". Most of these were on Census work but some of them were for the U. S. Air Force, the Army Map Service or the Atomic Energy Commission. We call these jobs "small" not because they were small in importance but because they were small in terms of the amount of Univac time they required as compared with the two "large" jobs our Univac performed during the period we operated it.

One of the two large jobs was the preparation of certain population statistics and the other was the preparation of certain housing-family statistics. Both were part of the 1950 Census of Population and Housing. Important characteristics of each of these large jobs were (1) they involved a very large amount of input (a total of about 20 million input items for both jobs combined) (2) relatively little processing of each input item was necessary (a fraction of a second was the time required typically for the Univac to dispose of an input item) and (3) a fairly large amount of output resulted—the output units were tabular presentations of population or family-housing statistics for a complex of geographic areas.

Commercial applications of large scale, integrated, high speed information processing equipment which come to mind most immediately are the activities under the jurisdiction of the accountants and bookkeepers. Whether these are more or less important than other possible commercial applications I am neither inclined, nor prepared, to argue. Control of all sorts of industrial processes will undoubtedly become more and more automatic through the use of these new electronic data processing devices. Such applications challenge the imagination and invite investigation and discussion. I feel competent only to conjecture and speculate about the impact of these equipments in this area.

The bookkeeping kinds of applications are, I believe, similar in many respects to the work we have done at the Bureau of the Census. Characteristics they share with Census applications are: (1) large numbers of input items (2) a small amount of processing per input item and (3) a significantly large amount of output. This is, of course, a generalization to which numerous exceptions can be found. However, it is, I think, obvious that maintaining inventory records on thousands of items, or preparing payrolls for thousands of employees, or sending monthly bills to thousands of customers or clients are applications much more akin to Census jobs than they are to the problems of engineering and applied mathematics to which large scale electronic computers have already contributed significantly and for which they were originally developed.

To our knowledge there exists nowhere, other than at the Census over eighteen months of continuous full time (around the clock, around the week) experience with high speed electronic information processing equipment on problems quite similar to those faced by the accountants in the business community. This does not qualify us as experts in such problems, ipso facto. In fact, the statement I can make with most conviction is that we have much to learn about how, most efficiently, to use our Univac. Nevertheless, we believe a few important comments are dictated by our experience.

The first of these is that large scale electronic information processing equipment can be more efficient for many commercial purposes than any other tool, or collection of tools, presently available. This statement is true, I believe, for devices which exist today, are commercially available and have a history of proven workability. Whether or not they can be used economically in any specific situation depends only on two things. One is the size of the job. At present the devices most readily available are very large scale and rather expensive. Obviously the job must be large enough to justify the investment in equipment. The other factor is the ability of the user to analyze and accurately define the final objectives of his present procedures; his willingness to arrive at those objectives by what may prove to be radically different procedures; and his willingness to familiarize himself with the logic of, and programming for, integrated information processing equipment.

With respect to the first of these two points I am confident that it will become decreasingly important. Equipments of reduced speeds and memory capacities but which are general purpose and well integrated are beginning to become available at costs significantly below those charged for their larger and more ambitious forerunners. This is one way equipment investment may be reduced. Another may well be through pooling of resources to defray the cost of equipment. This obviously may occur in several ways. A group of users might own an information processing system jointly; a trade association might acquire one to supply service to its members; an owner might sell excess time

on his system to other users; or independently owned facilities may sell information processing services. A wonderful attribute of these general purpose devices is the speed and ease with which they can be made to stop work on one problem and begin work on another. It takes literally only a minute or two to change. Last summer there was a period during which Census personnel were operating two Univacs at Remington Rand's plant in Philadelphia for three agencies. There were many actual instances where one minute a Univac would complete the numerical solution of a complicated formula for the Atomic Energy Commission and a minute later it would be tabulating Census statistics, and within a minute after it completed the Census work it would be engaged on a conversion of coordinates task for the Army Map Service.

This versatility means that to share an installation of electronic data processing equipment a group of potential users need not have common problems. Each potential user, however, must meet the second requirement I mentioned earlier. He must analyze and accurately define his end objectives and then develop procedures to achieve those objectives which take full advantage of the capacities of high speed data processing equipment. Not only Census experience supports this point. I have some familiarity with several investigations of the commercial applicability of these new equipments which were conducted by potential users. Two cases, in particular, I think bear mention. Both are multi-million dollar corporations with long histories of successful operation. Their businesses are quite different yet their investigation led to similar conclusions. These were that through radical changes in procedures significant economies could be effected through the use of these new equipments. In general these changes require the consolidation, into one integrated operation, of activities which are now being conducted in separate departments. In fact, in both cases most of the success of the investigation resulted from the fact that the investigaors were staff officers attached to the very top officials in their respective companies and therefore were able to cross existing departmental lines with little or no difficulty.

To restate, briefly, the first point concerning commercial applicability - equipment in operation and commercially available can be used profitably by the business community today if users will apply their detailed knowledge of their problems to the development of procedures to exploit these equipments.

Please keep the foregoing in mind. A person who believes he has perfect facilities for performing any task is either a genius or a fool. We at Census know we are not the former and hope we are not the latter. We are not satisfied with the equipment we have. I will try to indicate some ways in which we think it could be better. Unfortunately there have been instances where our indications of shortcomings have been interpreted as general condemnation of these equipments. This is completely unjustified. We are only trying to indicate how good equipment can be made better.

The input-output facilities presently available are badly out of balance with the ability of these equipments to manipulate information internally. There are two reasons for this. The first is a historical reason. These equipments were originally developed for applications where large quantities of input and output did not exist. Only when the designers and builders looked for commercial application did these facilities become important. The second reason for the mismatch between input-output and internal processing speeds is an engineering one. The input-output devices must make the transition from the physical world we humans know and operate in where time is measured in months, days and hours to the electronic world of the data processor where time is measured in seconds, milliseconds, and microseconds. This transition can be accomplished only by equipment with mechanical as well as electronic properties. We do not believe that designers have developed the best mechanical components for these equipments yet.

On the input side particularly, however, we think it is hopeless to expect mechanical elements to operate at speeds comparable to the internal speeds of the information processing equipment. This being the case we believe we must look for means to minimize the inefficiencies that slow input causes.

Procedures we used in the 1950 Census of Population and Housing will illustrate this problem. A census enumerator recorded the required information on a schedule, later a census clerk converted some of the information from descriptive words into number codes, still later a key punch machine operator transferred the intelligence recorded on the schedule to a punched card, and later still a card-to-tape machine transferred the intelligence from the punched card to the magnetic tape which is the input medium for our Univac. Now let us look at rates of speed with which these processes were accomplished. The enumerator's job involved much more than recording answers on the schedule. Let us say that the 30 or so households for which he obtained information each day was reasonable and satisfactory. The coding clerk processed the information for about 300 households per day. The key punch operator prepared cards for about 250 households per day, the card-to-tape prepared tape for about 10,000 households per day. The Univac processed information at a rate of about 20,000 households per day (all of this is based on an eight hour day). Obviously, the more manual the process, the slower it is. An ideal solution for this Census problem would be a device which would read the information recorded by the enumerator and transmit it directly to the electronic information processor thereby eliminating the coding clerk, the key punch operator and the card-to-tape operation and even the magnetic tape.

It is unlikely that we will ever achieve this ideal. We believe we are making progress toward it, however.

The National Bureau of Standards has been helping us with this problem and has in the final stages of development a device designed to read a census

schedule and transfer the information thereon directly to magnetic tape. In order to use this equipment we must require our enumerators to record their answers by means of positioned marks rather than descriptive words or numbers but this does not, at present, appear to us to be particularly burdensome or undesirable. We hope through the use of this device to affect substantial savings of time and money by eliminating the key punch operation and the card-to-tape operation. (There may be a partially offsetting increase in the time and cost of coding).

This we think illustrates the most important way in which input can be brought more nearly in balance with the information processing ability of these devices, namely by more complete integration of the processes involved. Possible commercial applications are very easy to visualize. For example: (1) cash registers which communicate directly with magnetic recording media very likely will be developed before long (2) the charge-a-plates now used in most department stores may well be modified so they initiate a communication with an information storage reservoir (3) standardized type used to record serial numbers and amounts on bank checks can probably be read by equipment which exists today. Other illustrations are not hard to find. In fact some of you may be familiar with the perforated garment tags which are currently being used by some mail order houses and department stores to mechanize the process of input of information to inventory control systems.

The mismatch between output speeds and information processing ability may be more serious for some commercial applications than it is for Census work. For some commercial purposes - account billing for example - there may be a one-to-one ratio of input to output items. Here high speed, legible printing is extremely important. It would be untrue and unfair to suggest that this problem has not received the attention and interest of designers of these equipments.

In this area, too, the Census has sought the assistance of the National Bureau of Standards. It is, I believe, correct to say that so far we have been advised to save our money. In other words there has not yet appeared output equipment enough better than our Uniprinters which type ten characters per second, to justify the price quoted for the faster equipment.

In summary of the second point then: There is room and need for improvement in input-output equipment and on the input side particularly there is need for system development and integration.

In conclusion, I want to remind you of the history of the development of punched card equipment. We, at the Census, are proud to have fostered Dr. Hollerith who invented the punched card method to increase the efficiency of Census tabulations. We recognize that this method grew to be the powerful aid to business and industry it is today as a result of the interest in it and the demands made of it by <u>private</u> business. Electronic information processing equipment which exists today was developed to meet government needs. Private uses are just beginning to appear. As businessmen put this new equipment to work for their purposes they will learn where it is strongest and where it is weakest. Just as their father's influenced the course of development of punched card equipment, today's businessmen should direct the development of tomorrow's electronic data processing equipment.