Problems of Life Assurance data processing

By Philip Giles*

The organizational problems that arise in Life Assurance data processing are considered in relation to the hardware and software employed in their solution.

Progress is steadily being made in the application of computers to commercial processes, but it naturally tends to seem least in those fields where the ultimate benefits are greatest. In finance there are possibilities of integration often not present in industry, since the major output is generally printed information which can be directly produced from a computer. In Life Assurance these advantages are enhanced by the facility for rapid arithmetic. The long term nature of Life Assurance contracts, and the volume of mathematical and statistical knowledge required for the successful transaction of this type of business, indicate that extensive calculation is necessary in order to examine the relationship between liabilities and assets. Only by this examination can profitability be judged and future policies decided.

This paper considers some of the alternatives presenting themselves to the prospective user of a computer for Life Assurance processing who is aiming at an integrated main record file updated daily. Any conclusions are necessarily entirely personal opinions, but it is hoped that they will stimulate critical thought on the points in question. In any particular situation there is probably a best solution, but circumstances vary so much between different workers in this field that final systems are unlikely to bear more than a slight resemblance to each other. In particular no mention is made here of the problem of maintaining branch or agency accounts. These have been discussed in the Reports of Study Groups B and E of the Faculty of Actuaries (1962) where the data-processing requirements are analyzed.

The problems and progress of the large insurance companies in the United States were thoroughly reviewed by Finelli (1960), but most insurance organizations in Great Britain form smaller units where the economies of computer usage are less easy to demonstrate. Sutton (1960, 1963) has reviewed 5 years experience of an office with 365,000 policies, and Engelfriet (1960) has discussed the problems encountered with 100,000 policies. These two sizes span the range considered in the following sections.

Program storage and organization

The programmer's ideal is a single-level store, but experience shows that, as programs grow, the immediate-access store becomes too small, and magnetic-tape library techniques are employed for part of the program.

Programs fall into natural sections when being written. Each section should be independent of all other sections, to facilitate future alteration of a part of the system. In large programs each section is tested and debugged separately. Sections may be sequential or possibly in parallel, but in any case are best checked before amalgamation. Sections will vary in size but may be, say, 50 to 1,000 instructions in length. Some sections will be sequential, and hence the order in which they are required is known. This permits efficient storage in one dimension, as on magnetic tape. Some sections, however, will be on alternative parallel paths in the flow chart, and may be required on few occasions. If these are stored sequentially they will slow up access to the next required section. Some other sections will provide general-purpose routines which may be required by the main program at many different points—nearly necessitating true random access. These are most inefficiently stored in a magnetic tape library.

On the other hand, if a backing store is available in the form of a magnetic drum with a reasonable capacity in each track—somewhere between 50 and 1,000 instructions and nearer the latter—then all sections of program are available when required and generally at a transfer rate at least as great as that from magnetic tape. If each track contains one section it is possible for hardware to commence transfer at the next available drum word and to add that word number on to the destination address. When the nominal first word of the section is reached on the drum the fast store address jumps back to the beginning of the section. This system cuts mean access time down from half a drum revolution to half a word transfer. The moving parts in the mechanism are fewer than those in a tape transport, and greater mechanical reliability can thus be expected. On price the cheaper magnetic-tape transports cost about the same as a magnetic drum with a capacity of 10,000 to 20,000 instructions, and in most applications one tape transport would have to be tied up on program library work. At this cost (between £5,000 and £10,000) the transfer speed is much higher from the drum than from the tape. For larger programs the comparison is between a larger drum (30,000 instructions) and one tape transport working at a similar fast transfer speed. On this basis tape has little if any advantage over drums. The availability of program sections in random order from a drum generally means that less fast storage is needed, and this can save a considerable part of the cost of a machine—fast storage is between 10 and 50 times as expensive as storage on a magnetic drum.

No doubt these are some of the considerations that led to the incorporation of drums as backing stores in

* Scottish Amicable Life Assurance Society, Craigforth House, Stirling.
Atlas I (see Howarth et al., 1962) together with a supervisory program to enable the programmer to work as though there was only one large single-level store. Such considerations apply in Life Assurance administration because of the enormous size of programs required in any integrated system—between 25,000 and 100,000 instructions is the normal experience of those working in this field.

Input and output

Due to the low activity of Life Assurance records, one of the major problems of organization in this application of computers is the simultaneous handling of a variety of different forms of input and output. Centralization of premium collection—by banker’s order where possible—simplifies the major problems of input. The accuracy of this input can be further improved by the use of pre-printed credit transfer forms transmitted to the remitting bank in advance of the payment date. This system has already been applied to monthly payments under hire purchase contracts. The forms should be printed on cards punched with the amount and serial number.

For output very high speed is not nearly so important as flexibility of format, and of stationery. This is particularly true if daily updating, with its many advantages, is to be a major aim of the system. This immediately leads to considerations of buffers for temporary storage, followed later by off-line printing. The computer itself can quite reasonably be used for those forms of output that require a large volume of printing, but many forms of output are used so little that the punching of cards or paper tape followed by printing on much slower equipment is more economic. This is so because the time wasted during the setting up of a slow machine by an operator is much less valuable than the time wasted during the resetting of a high-speed line printer. No one has yet produced a printer that can print from magnetic tape at a price as low as that available for paper tape.

The buffer problem can be reduced by breaking down the complete file run into sections—one for each magnetic tape used—and taking advantage of the tape rewind time to clear the buffers of output and refill them with input. Few small computers can handle input and output at the same time as magnetic tape reading and writing without either slowing the latter down or producing an extremely complex program. It is thus preferable to separate the work into its two phases—clearing the input/output buffers, and searching the file and updating it. If these phases change over fairly frequently the buffer does not need to be very large, and part of a magnetic drum may well be sufficient. This avoids the use of tape transports for buffers and, by using up the tape rewind time, makes it unnecessary to have additional tape transports to permit a queueing system for the main file. The queueing system is normally employed to ensure that two further reels of main file magnetic tape are working while the previous old and new reels are being rewound.

This system allows the use of comparatively slow tape for the main file because the buffers are working on a drum of high transfer speed. These buffers will have sufficient time to print and punch out at full speed while the main tapes are rewinding. A card punch can handle several formats at once because they can be subsequently separated on a card sorter before listing, but other outputs can usually only handle one format.

While tape reading and writing is in progress the program controller may well be under-utilized. If this is so, it can be instructed to apply any consistency checks to each record for which the necessary redundant information has been recorded, and to carry out a statistical analysis of the policies contained in the block currently being written to tape. The results on any one run should be gathered into a reasonable number of pigeon-holes. Any really extensive analysis should be broken down into sections for running on consecutive days. The appropriate keys for each section will be inserted into the analysis program before the daily run.

Hardware for the main record file

Life Assurance records can be looked upon as basically a form of account keeping, though the ancillary records required in addition to the cash position are fairly extensive. As a result the policy record will extend to between 250 and 500 characters in length. The variation in type and complexity of individual policies of assurance makes a variable-length record essential; otherwise more than half the reading space on the main file would be occupied by voids. Many policies only have one benefit—the sum payable on death being fixed—and only one address—that of the contributor, who is also the life assured and beneficiary; but some policies have two or three separate benefits and sometimes more than one address. If routine work for all policies is to be carried out on the computer, space for all possible contingencies must be provided.

At first glance it would seem reasonable to keep the accounts on a cash basis, and to ascertain outstanding premiums by an examination of the due date only when they become sufficiently overdue to require action, or on an accounting balance date. In practice, however, the rate of premium may alter before the last premium at the old rate has been received, and administration is much simplified by performing a debiting action and recording the outstanding premium as a debit item until it is received. This means that a continuous balance can be maintained on the accounts from day to day, and Balance Sheet figures are immediately available.

It can now be seen that updating of the file consists of the following main categories:

(a) Implicit updating
   (i) Creation of an outstanding premium entry on the due day.
   (ii) Creation of Premium Renewal or Reminder Notices when required.
(iii) Intimation of a change in the conditions of the contract such as expiry of premium payments or maturity of an endowment assurance.

(b) **Explicit updating by serial number**

(i) Receipt of a premium and cancellation of the outstanding entry.

(ii) Receipt of a simple change to the account, such as change of address.

(iii) Request for surrender or alteration of the contract.

(c) **Explicit updating by some other secondary key**

Examples of this requirement are the changes required in a large number of records when a new servicing branch is created, or when two agencies are amalgamated.

If the file of policy records consists of an equal number of monthly and annual policies then the activity level can easily be calculated. This composition of the main file is, of course, much simpler than a typical composition, since quarterly and half-yearly policies are usually present, but it is a distribution of frequency of premium payments towards which many Life Assurance offices are moving. With this composition there will be 12 × 3 + 1 premium debits and a similar number of premium receipts each year. This gives an activity of 5% per working day, to which must be added an allowance for premium renewal notices, changes of address and many other minor types of alteration. Since renewal notices are not normally issued for monthly premium policies, the total activity will be about 6% on these assumptions. In practice, in many offices at present the volume of monthly business is lower, and 4% would be a reasonable figure.

The second form of updating and the low activity level suggests the use of random-access files held in policy number order, but the first form suggests a file in order of due date. Ideally the file should be held on a two-dimensional lattice system providing access immediately by number, but grouping also by due date. It is essential that the records can be updated fairly promptly, and with the increase of monthly premium business there is a lot to be said for daily updating. Thus a form of record with a fast transfer speed is required so that a search can be made throughout the full records for those accounts to be updated. This can be achieved with magnetic discs or drums, but for the extent of storage required—50 to 100 million characters—a form of storage physically limited in capacity is rather expensive (perhaps 400 characters per £). However, a demountable system with an access time of about a second is now available from a number of manufacturers, and provides the facility of interrogation on demand to about 10,000 policy records, which makes immediate replies to inquiries possible. Nevertheless, it is doubtful whether this leads to quicker file updating than that obtained with magnetic tape if only 4% of the records require alteration, because the time to obtain access to one record on a random-access system is as high as the time to copy 24 unwanted records on tape. The random-access disc transport will probably cost more than a tape transport.

The alternative is to use magnetic tape. The cost of using tape is much lower because the medium of record is no longer a physical part of the machine requiring for every record a machine address. It is now information that is read serially at one machine address—the tape transport. Since the length of a policy record may be increased by an alteration, a complete fresh tape must be written either at the same time at another tape transport or at a later time, having the correct records properly filed among those which remained unaltered.

The frequency of updating is limited by the time taken to read and write the whole file. Assuming a file of 200,000 policies this amounts to nearly one hour for 30 kc/sec magnetic tape if both are performed simultaneously, or 2 hours otherwise. Since each tape transport costs about £4,000 plus one-third of the speed (in cycles per second) the cost of two transports at 30 kc/sec— including an allowance for the control unit—would be about £28,000 and this could deal with the file of 50 to 100 million characters on a daily system, provided that arithmetic speed was reasonably fast.

If magnetic tape is used, a method of answering inquiries as to the details of a given record must be supplied. The following day's normal tape run can be used for the purpose, but it is possible to save the delay by printing out every altered record, together with a daily index to these detailed printed records in the format of a double-entry table. This can, on these levels of activity, be restricted to a reasonable size provided that the file does not exceed a few hundred thousand policies. At the same time this complete record of file alterations provides a historical record of all changes to any one policy. If each print includes the date of the previous print for that policy then historical inquiries can easily be answered, and the auditor can trace any individual transactions and their effect on the main records, although these records, being on magnetic tape, have since been altered several times.

**Format of main records**

Access is required to any record on the main file either to make an explicit or an implicit alteration in that record. Thus it is essential to be able to search the file—which is recorded in variable-length format—and to extract the portion which requires alteration. It is always easier to compile programs which handle fixed-length records with specific information always in the same position in the record, so it would appear that a prime requisite of this system is a subroutine that will extract from the complete file only those sections which require action, and will convert them from variable to fixed length. This is particularly so if internal transfer speed is high (e.g. 400 kc/sec) and tape reading comparatively slow (e.g. 30 kc/sec). Since most items of updating only require a section of a policy record, the block of records on magnetic tape should be divided into two levels—policies and sections of policies.

During a file search many records will be unwanted,
Life Assurance

and the writing of the new file can be speeded up by inserting a master key in front of every block of policy records. This will contain the highest serial number in the block and the earliest implicit alteration date—thus permitting an immediate decision to be made to write away the block unaltered if this is appropriate. If not, then the sections must be individually examined to find those requiring action.

Since action may necessitate a change in the length of the block of records, each section must be written to another area of fast storage before the next section is examined. For these purposes it is essential that each section carries a record of its own length. It would be attractive to have internal cross-reference between sections such as those available in list processing techniques—see Woodward and Jenkins (1961)—but any alteration in the section length would require an alteration in the cross-references since storage is scarce and voids must be minimized. The resulting complications in programming updating procedures would probably outweigh the advantages in moving quickly from section to section within a policy record.

In order to overcome the lack of internal cross-references the location of the beginning of the first section of a policy record should be retained until the last section of that record has been tested and transferred to the write-to-tape area of fast storage.

The first section should contain the serial number by which explicit alterations are identified, and subsequent sections must contain keys identifying them so that the program can find them by searching within the policy record. Sections liable to implicit alteration should, where possible, be placed behind those liable to explicit alteration, so that both may take place on the same occasion without confusion. Nevertheless, it should be borne in mind that the delay of an alteration for one day is usually quite acceptable, and may make for simpler organization, since the possibility of coincidence can thus be avoided.

This arrangement of records should permit a layout of the file that provides two areas for a full block of data from tape in variable-length format—one for reading from and the other for writing to tape—and a third area for one section only of a policy record in fixed-length format. The first two areas are essential if simultaneous read and write are to be employed. The third area is economical in space when compared to that required for the largest policy record in fixed-length format, which is the obvious alternative. This alternative would require subroutines for searching and extracting inside records which would be of similar complexity to those for searching and extracting inside sections of records.

These considerations would appear to favour the use of a machine based on character addresses in preference to one based on fixed word-length addresses in the organization of fast storage and program instructions. Against them must, however, be set two other considerations. Firstly variable-length organization of the hardware of a machine is most suited to alphabetical data which forms no more than about a quarter of the policy record—lengthy policy clauses being excluded except in code form. The balance of the record consists of numerical data which is more efficiently recorded in 4-bit characters rather than 6-bit. Secondly the arithmetic speed of a fixed word-length machine is much better than that of a character-addressable machine of the same generation, because all the digits can be handled in parallel by the mill. There are a few essential operations in Life Assurance that require a large volume of arithmetic even if the records are pre-sorted, a time-consuming operation, and the number of such operations can be expected to increase as facilities for them are improved.

The future

These considerations suggest that it should be practicable for a Life Assurance office with two or three hundred thousand policies on its books to operate a largely integrated system with daily updating at a cost of about 10/- per policy record for the computer alone, with a minimum of £125,000. This should provide a greatly improved service both to its customers and to its own management. In the past some requirements in the form of statistical and mathematical analysis have been restricted by the time and labour required to produce them. This need no longer be so, and the challenge to management will become one of formulation of requirements.

References


