Patent Office
Research and Development
Reports . . . . . No. 6

INTERRELATED LOGIC ACCUMULATING SCANNER (ILAS)

Prepared by

Don D. Andrews
Director

Office of Research and Development
U. S. Patent Office

June 25, 1957

Robert C. Watson
Commissioner of Patents

Sinclair Weeks
Secretary of Commerce

INTERRELATED LOGIC ACCUMULATING SCANNER (ILAS)

It seems probable in these days when the trend is so unmistakably toward larger and larger and increasingly faster electronic computers and data processors that many of us may have lost sight of the numerous advantages of punched cards and punched card techniques without having expended greater effort to overcome their limitations.

Many experts recognize the comparatively low cost, high reliability and simplicity of operation of punched card equipment yet turn to expensive, complex digital computers principally because of the limited number of bits of information which can be recorded on a single card and the low rate at which these cards can be fed through the equipment. These punched card factors may be considered as limitations as to their use but it is suggested that more efforts could be profitably expended in removing such limitations rather than relegating punched card techniques to a "still-useful-but-obsolete" status.

The Patent Office Research and Development group for the past two years has devised several systems for coding the technical and scientific content of patents or other literature in such form that the codes can be mechanically processed to select or identify only those patents having the required legal pertinency thereby providing a basis for allowing or not allowing applications for patent. In devising these codes the Patent Office found that the high level of detail needed and the wide scope of subject matter involved in patent work necessitated the use of relatively long complex codes.

While the Patent Office has had access to the National Bureau of Standards electronic computer, SEAC, it still felt somewhat handicapped because many features of logical design in the coding system could not be tested quickly or easily due to the facts that the necessity of first preparing and "debugging" a program and the task of either learning the techniques of programming a computer or explaining the ramifications of the logic to be tested to an expert programmer made such tests expensive and time consuming.

The Patent Office staff, therefore, turned to punched card techniques to provide such a research tool and, of course, was confronted with the limitation of 960 bits of information on a single card that could be correlated together in a logical relationship. The problem of speed was not thought to be too serious in view of the research nature of the work, the probable gradual build-up of the file of punched cards and the likelihood that the total file could be subdivided on a basis that would require only a portion of the file to be processed for most operations.

This problem of space limitations on the card was partly solved by Mr. L. A. Wilson* while working with the Patent Office staff back in 1950. This system was described in a paper presented before the American Chemical Society in September of 1951 and later published.[(1)]

The method of coding in this system necessitated a quite unconventional arrangement of punches on the cards as well as several minor unspecified alterations in the operation of the Electronic Statistical Machine, Type 101, of IBM in order that it could process the cards.

Since the description of this system of coding has not received wide circulation a few words of explanation of it might be in order here.

In the IBM card as conventionally punched each of the 80 vertical columns may have any one or certain combinations of two of the twelve rows punched to represent any one letter or numeral. Since the card passes under

*Manager, Card Machine Development, International Business Machines, Endicott, N. Y.

the reading brushes of the 101 machine while moving in a direction parallel to the vertical columns the presence or absence of any hole or combination of holes in each of these 80 columns can be determined only after all 12 horizontal rows of the card have passed the reading brushes and effected the transferred or non-transferred status of a series of recode selectors or relays. After sensing the last row of punches in the card and before the feeding of the next card the relays are electrically tested for conformance with a preselected pattern of holes and if agreement is found the card is deflected into one of a series of sort pockets rather than into the reject pocket. All relays are then restored to their original condition prior to the reading of the next card.

In the 1950 system of coding, holes or combinations of holes which denoted a single numeric character were all punched in the same horizontal row on the card rather than in the same vertical column, as formerly. Since there were as many as 80 binary bits available in each row a whole series of numbers were represented in each row which thereby formed a code group or, in computer terminology, a "word." Since there were twelve such rows, as many as twelve words could be contained on a single card.

The changes in the 101 machine, so far as the author knows, have never been published but, in operation, the machine tested each row of holes for a desired pattern before the next succeeding row of holes reached the reading brushes. If such test were affirmative the transfer of a relay took place and its transferred state continued for all succeeding holes of the card as well as for an indefinite number of succeeding cards until such time as the relay ultimately was released by a specially located control punch in one of the cards which would then cause the sorting of the last card.

By this means twelve, rather than one, logical decisions were made for each card but what is most important is that coded data of any length could be processed as a single record without being confined to a single card.

In June 1956 it was decided that the needs of the Patent Office required that it explore further the possibilities of punched card techniques with the chance of procuring an improved machine having advanced features of the modified IBM 101. The Patent Office turned to another agency of the Commerce Department, the Census Bureau, for assistance. This agency had previously designed and constructed for its own use a number of type 488 Multi-Column Sorters which had many features in common with the commercial version of the IBM 101, including simultaneous reading of all 80 punch positions in each horizontal row.

Armed only with the knowledge of the general operation of the modified IBM 101 and certain additional characteristics which would be desirable, work started on the modification of the Census Multi-Column Sorter. This was the beginning of ILAS.

All equipment was removed from the sorter except the card handling and the card reading brushes. A new console was built to house the logic circuitry having its input coming from the reading brushes and its output going to the card sorting magnets of the sorter unit through an interconnecting cable.

Fig. 1 shows a view of the complete ILAS machine.

Column Relays

Eighty relays were allocated to the function of detecting the presence or absence of a hole under each of the card reading brushes- twelve times during each card cycle. Since it was desirable to retain the normal card feeding rate of 400 to 450 cards per minute each row of holes must be read in something less than 8 milliseconds. This means that these relays must pick up, pass a test pulse through their contacts and drop out before the next row of holes is read 8 milliseconds later. IBM relays were found to have satisfactory characteristics when energized through 1D21 thyratrons. The column relays were dropped out by opening the plate circuits of the 1D21's with a bank of four 4-contact (mercury-wetted) relays controlled by a timing pulse from the sorter.

Rotary Switches

Since each of the column relays has twelve independent sets of contacts having normal and transferred positions it is seen that 12 different combinations of 80 punches may be detected by proper interconnection of all corresponding sets of contacts on each of the 80 relays. To make these interconnections through plugboard wiring would be difficult and almost impossible to recheck to find any error in wiring, consequently a series of rotary switches was mounted on the front panel of the console for making these interconnections. Each rotary switch had 20 positions on each of four levels and made the interconnections between common, normal and transfer contacts of corresponding contact sets of 4 successive column relays. Each rotary switch was thus able to be preset for all of the 16 combinations of hole patterns according to the hexadecimal notation, (i.e. 0 - 9 and A - F) but in addition one position (I) served as a shunt for all four column relays with which it was associated and another position (H) served as a shunt for these as well as all the remaining column relays of the row which are under the control of the switches.

Punched Card Arrangement

Since this machine was to be used on research problems it was most desirable that it be designed for the maximum possible flexibility yet incorporate certain advanced features for correlating the recorded data. To this end the punched card was laid out as shown in Fig. 2.

The first four columns of each horizontal row provide a "Signal" field in which is punched any one of 12 different combinations of holes to be used as markers. For example if each row of holes in the card is considered to be a "word" consisting of a number of different hexadecimal characters, then a group of such words separated by a distinctive signal could be a "phrase", a number of phrases could be grouped to make a "sentence" and several "sentences" grouped to form a "paragraph." A distinctive signal would also be used to mark the end of all the codes representing each patent or document. These signals serve very much the same logical function that parenthesis and brackets serve in mathematical expressions.

The first four column relays have corresponding sets of contacts permanently interconnected so as to respond to 12 different signal codes (1-9, A-C) with their outputs brought out to the plugboard so as to be available for picking up, dropping out and testing other relay circuits. In addition those contacts not used to set up each signal code are interconnected and brought out to the plugboard to supply the negation of that signal, that is, there will always be a pulse present whenever any signal is read other than that particular signal.

Modulant

Columns 5-8 of the card is the Modulant field. This single hexadecimal character is used to modify the meaning of the remainder of the code word. The first rotary switches of each of the twelve groups permits the selection of any modulant code while plugboard wiring is used to make the connections to inputs and outputs of the moduland column relays.

Subject Matter

Columns 9 - 68 of the punched card contain the basic identifying codes for the information being processed. These are divided in the code word into 15 hexadecimal characters of four bits each. Column relay contacts are interconnected by the rotary switches to set up a maximum of 12 different question words having up to 15 hexadecimal characters each. By this means it is possible to distinguish 2^{60} different code combinations for each code word within the Subject Matter field each of which can be in combination with the same number of other code word combinations. However, not all 15 characters need be used in every word, for the shunting positions of the rotary switches on the control console allow the machine to disregard any individual or series of characters.

The flexibility of the machine is greatly enhanced by the facility for altering the mode of comparing the question words set up on the rotary switches of the control console with the code words punched in the cards. Normally the comparison is on a bit by bit basis for each character selected for comparison and results in an exact pattern match. The basis of comparison may be shifted at will on a whole word basis from this exact pattern match to a "set-subset" relationship. That is, if a hexadecimal character such as "B" (X-XX) is punched in a card, question characters such as "1" (- - - X), "3" (- - X X), "8" (X - - -) as well as "B" will be satisfied by the punching in the card. This is very important when dealing with Patent Office subject matter for the fact that a patent disclosure contains the sought for items is ordinarily sufficient notwithstanding the presence of additional disclosure. A simple example may make this clear. If a single four bit character is used to represent the different kinds of chemical elements present in a chemical ring structure the first bit may represent carbon, the second oxygen, the third nitrogen and the last sulfur. If all four elements are present in a single chemical ring, all four bits would be punched yet that character would respond even though only the first two bits, representing carbon and oxygen, were requested by a setting of the rotary switches to "C".

The shifting from the exact pattern to the set-subset comparison is effected by twelve manually thrown toggle switches at the left of each horizontal row of rotary switches on the control console, each controlling a bank of five 12-contact relays.

Interfix

The modulant and subject matter fields in the code word serve very flexibly as a means for identifying the various segmental units which collectively represent the total disclosure of a patent or other technical document but for patent problems this is seldom enough. Almost any piece of machinery or equipment when reduced to its component parts has lost its identity and individuality and differs only in an insignificant manner from a quite different piece of machinery similarly reduced to components. A television set when reduced to a list of component parts likely would be indistinguishable from a list of components for a radar set.

That which distinguishes a radar from a television is the relationship which is present between each of the components and every other component. Therefore, it is most essential that these relationships be identified in coded data representing the subject matter of a patent.

The signal character at the beginning of each word serves to group a number of code words together and this grouping can be understood to mean that each of the components represented by the individual code words of a group bears a certain specified functional, structural or electrical relation to each of the other components of the group.

This grouping is most useful but it is still not flexible enough, for many components must be represented logically as being in two or more different groups. A code representing such a component could not be physically located in more than one group and repeating the code for each group would lose the relationship of the several codes as pertaining to but a single component.

An approach to a solution of this problem has been incorporated in ILAS. It makes use of a technique which we call "interfixing". By this technique one incorporates into a code word representing a component part of a device one or more sets of additional distinctive marks, each set representing a different relationship among the component parts and each mark of the set representing one of several such different relationships.

A simple example may help make this clear. A gear identified as "A" drives shaft "B", and shaft "B" in turn drives pulley "C". If we wish to interfix the driver-driven relationship of these parts we affix to the code for A any one of a set of marks, such as "3" and add the same number to the code for B meaning that the driver-driven relationship is present if the two codes contain the same, although unspecified, interfix number. Similarly, B would be interfixed to C but a different arbitrary interfix number would be selected, such as "5". The interfixed codes would then look like this:

A - 3

B - 3, 5

C - 5

Note that "B" participates to two driving relationships.

Columns 69-80 of the ILAS code word is reserved for recording 12 different interfix markings. This card area may be handled either as a single field or as multiple fields of 2, 4 or 6 columns each. In any interfix field, punches in the same interfix column in rows of two or more code words mean that those code words have the specified relationship. It is arbitrary as to which column is selected to represent any individual relationship within a field. The ILAS machine has means hereafter described for discriminating between a series of codes which are interfixed from the same series which are not interfixed.

-Hit Relays-

In addition to the column relays, selected ones of which must pick up and drop out for each of the 12 rows of the card, "Hit" relays are provided which "remember" that a part of the specifications of the total search request has been detected and when all of the specifications have been met, even though many cards may intervene, instruct the sorter to alter the path of the card.

ILAS has 24 such hit relays each of which has separate pickup and drop out connections and common, normal and transfer contacts.

Forty eight interfix hit relays are interconnected to "remember" each punch in the 12 interfix columns for up to 4 different code words interfixed in the same column.

When a series of code words are to be found irrespective of any specified interfixed relationship the simple hit relays are picked up by code word signals getting through the column relays and subject matter rotary switch combinations but where a further specification of an interfixed relationship is needed the interfix hit relays are used.

Fig. 3 shows the arrangement of relays in ILAS without attempting to show the various possible interconnections among them.

-Plug Board-

The rotary switches of the control console are used to quickly and accurately set up the various code words of the search question or specification. On the other hand the logic or relationships required to be found to exist between code words is spelled out by appropriate wiring of a small fixed plug board. No interchangability was provided for this plug board because of the extremely remote possibility that the logic of any two search specifications would ever coincide.

The layout of the hubs on the board is such that as many as possible of the connections will be made by simple jack plugs. In this way the connections may be checked more easily than would be the case with bunched wire connections.

Each of the signals and their negations, the inputs and outputs to each of the modulant and subject matter fields are brought out to conveniently arranged hubs as are the pick ups, drop outs and contact connections for the numerous hit relays.

Timing

The operation of the machine is timed by a series of four closely spaced pulses (T_0, T_1, T_2 and T_3) each occurring 12 times for each card cycle. Pulse T_0 tests for the presence or absence of a hole under each of the reading brushes and picks up the corresponding column relays. T_1 tests the column relay contacts to pick up the hit relays and T_2 and T_3 are used to test and pick up all relays connected for higher orders of logic.

-Progressive Sorting-

The fact that any number of punched cards may be effectively coupled together in a complex logical relationship necessitates the maintenance of the punched cards in a precise ordered sequence, otherwise true relationships are lost and false ones established. This is all the more critical because all of the 80 columns are used for coding with none reserved for housekeeping information identifying individual cards.

The operation with the modified IBM 101 resulted in the last card of any sequence being diverted into a sort pocket when the codes of that sequence satisfied the search specification. This necessitated a manual card refiling operation after every run of the machine before the next run of the cards could be made.

The best solution of this problem, of course, would be a printer which would print out the document identification number from the information contained in the last row of holes punched in a sequence whenever the preceding codes have satisfied the search specifications. Such a printer, however, would be expensive and it alone probably would cost more than the remainder of the equipment. A compromise was arrived at in the use of "progressive" sorting.

By "progressive" sorting it is intended that the ordering of the card file is not disturbed when a sorting operation is called for as a result of the card data satisfying the search specification. Instead, the cards at the start of a run are fed into the reject pocket until the first sort is indicated. At this time the cards are deflected into the next adjacent pocket to which they continue to be fed until a second sort is required. Then the cards are again deflected to the next adjacent pocket and so on until all pockets are used. The bottom card of each pocket, other than the first, identifies a document selected by the machine. The ordered state of the card file is maintained by merely restacking the cards from each pocket in sequence. A manual switch on the sorter can be set to stop the machine whenever there are no more pockets for additional sort operations or alternatively to allow the machine to continue but then to start feeding cards again into the first pocket which meanwhile must have been emptied of cards.

-Speed-

The ILAS machine was designed to handle cards at the rate of 450 cards per minute but on occasion this has been increased to more than 500 cards per minute before timing difficulties resulted in erroneous sorting. As an aid to testing, provisions were made to operate the machine as slow as 45 cards per minute.

At the rate of 450 cards per minute 90 rows of holes or 7200 bits of information are scanned each second. This is approximately one-twelfth of the scanning rate of commonly used magnetive tape input units (90,000 bits per sec.) but the fact that 12 different sets of comparisons are made simultaneously on the fly without additional computation time further reduces the gap between punched card equipment and electronic computers costing at least 100 times as much as the ILAS machine.

It should also be noted that the time required to prepare and load the machine with the data defining the search specifications usually takes only a matter of minutes and is a direct operation subject to visual checking whereas preparation of data for a computer usually must pass through several hands and, if trouble develops from the data, requires elaborate checking procedures.

The magnitude of the entire Patent Office searching operation is such that ultimately the largest and fastest available computer or specialized machine will be necessary but during the present experimental phase and until our file of data increases beyond limits not yet fully determined, the Patent Office expects to continue to utilize punched card equipment.

Reference

(1) Mechanized Searching in the U. S. Patent Office. M. F. Bailey, B. E. Lanham and J. Leibowitz, Journal of the Patent Office Society, Vol. 35, pp. 566-587, August 1953.

FIG. 1

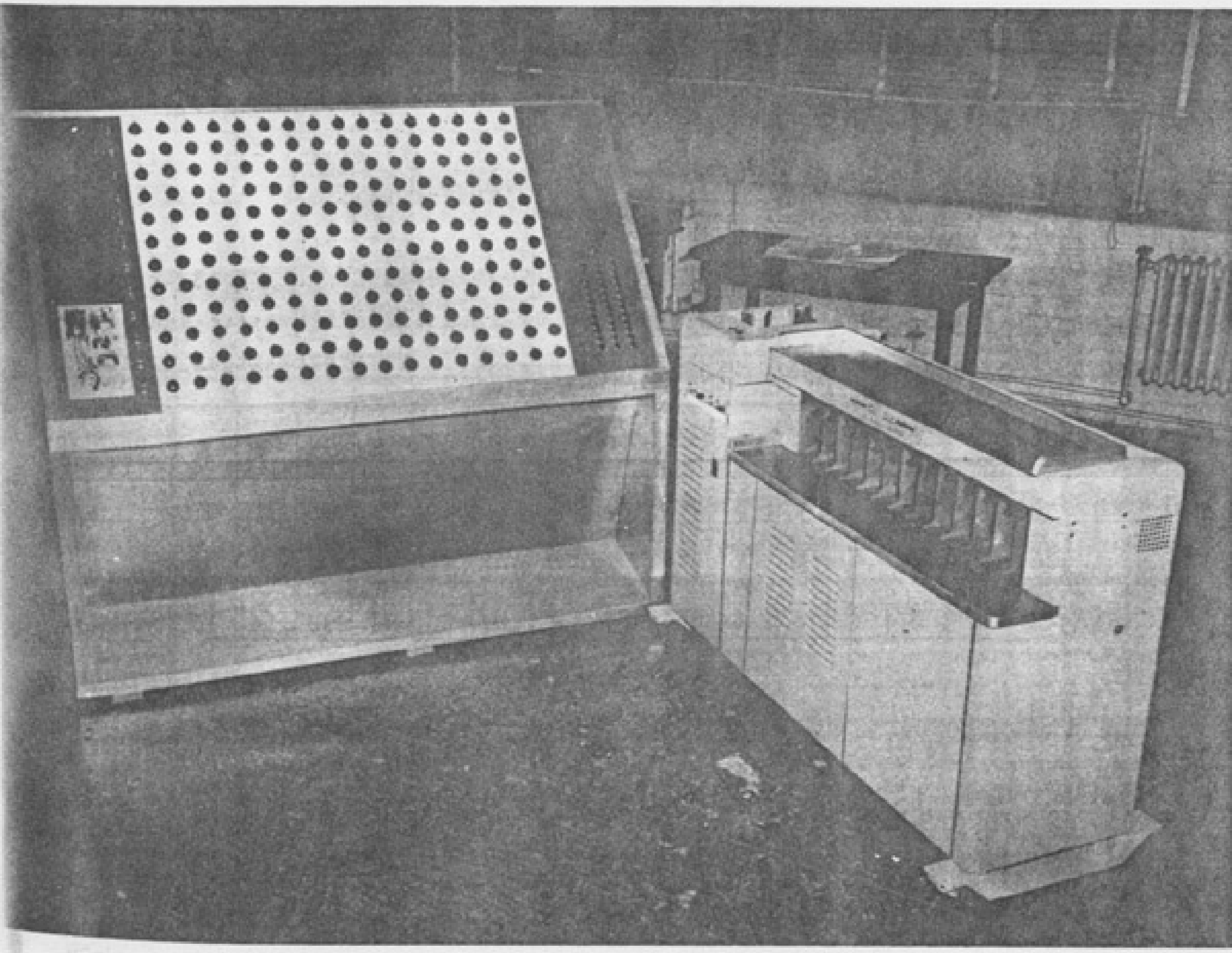
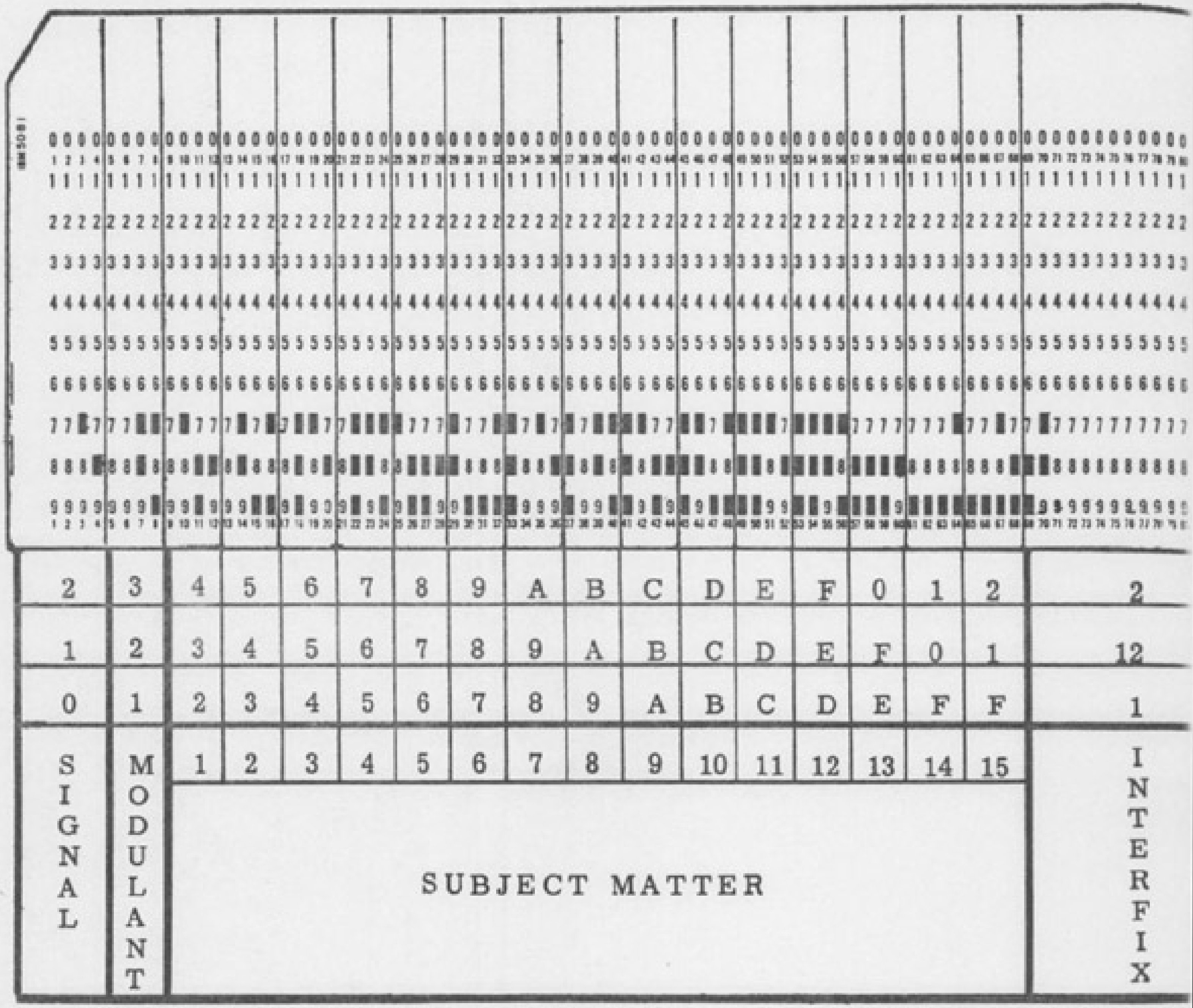
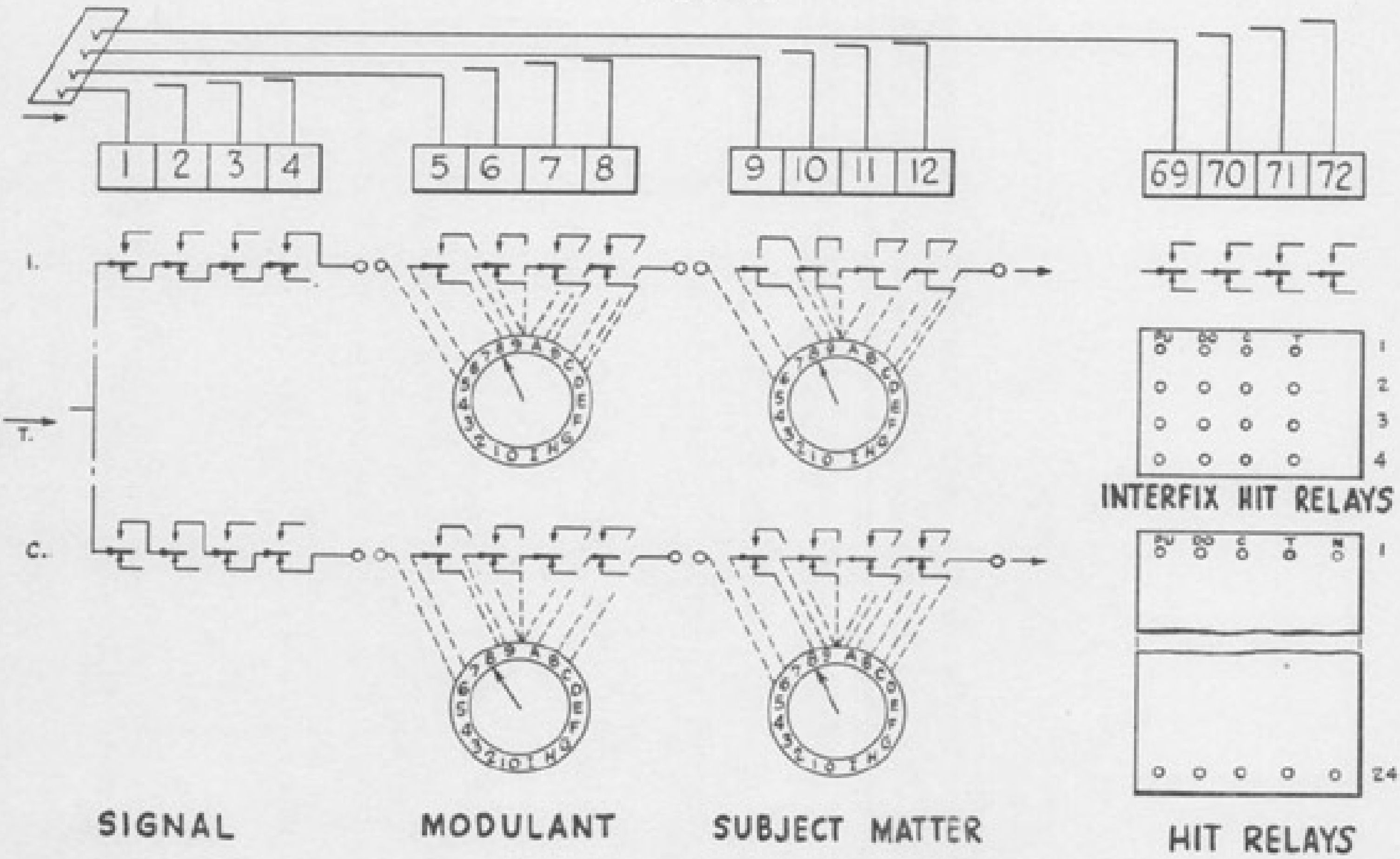


FIG. 2



2 3 4 5 6 7 8 9 A B C D E F 0 1 2 2
1 2 3 4 5 6 7 8 9 A B C D E F 0 1 12
0 1 2 3 4 5 6 7 8 9 A B C D E F F 1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
SIGNAL
MODULANT
SUBJECT MATTER
INTERFIX

FIG. 3



ILAS RELAY ARRANGEMENT