

A STUDY IN MACHINE-AIDED LEARNING

by

Chung Laung Liu

B.Sc., Taiwan Provincial Cheng Kung University, 1956

Submitted in Partial Fulfillment  
of the Requirements for the  
Degree of Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
June, 1960

Signature of Author Chung Laung Liu  
Department of Electrical Engineering

Certified by \_\_\_\_\_  
Thesis Supervisor

Accepted by \_\_\_\_\_  
Chairman, Departmental Committee  
on Graduate Students

## A STUDY IN MACHINE-AIDED LEARNING

by

Chung Laung Liu

Submitted to the Department of Electrical Engineering  
on May 21, 1960 in partial fulfillment of the requirements  
for the degree of Master of Science.

ABSTRACT

The conventional teaching machines are studied. The advantages and the possibility of using a computer as a teaching machine are investigated. A program using an IBM 704 computer to teach arithmetic and a program using an IBM 704 computer to teach matrix algebra are suggested.

Thesis Supervisor: Ronald A. Howard

Title: Assistant Professor of  
Electrical Engineering

### ACKNOWLEDGEMENTS

I would like to express my appreciation to Professor Ronald A. Howard who has initiated my interest in this field of research and whose constant encouragement and advice have made possible the completion of this thesis.

This work has been supported by a fellowship from the Li Foundation, Inc., to whom I express my deep gratitude.

I am indebted to the M.I.T. Computation Center for providing facilities which are indispensable to this research.

To Miss Theodate Coughlin credit goes for the typing of the manuscript.

TABLE OF CONTENTS

Chapter 1	Introduction	page 1
Chapter 2	Conventional Teaching Machine	4
Chapter 3	Computer as a Teaching Machine	16
Chapter 4	A Program Teaching Matrix Algebra	27
Chapter 5	A Program Teaching Arithmetic	36
Appendix A		46
Appendix B		50
Appendix C		59
Bibliography		63

CHAPTER 1INTRODUCTION

One of the characteristics which distinguishes the human being from lower animals is his ability to learn. The most fundamental way of learning is to learn through one's own experiences. One learns to keep his hand away from the fire because he has been burned by the fire before. This way of learning is primitive and ineffective, since it is impossible for someone to have sufficient time and opportunity to experience all the things he wants to learn about. Moreover, there are so many events which are not possible or advisable to experience, e.g., starvation to death. A great factor credited for the cultural evolution of human beings is that people began to learn not only through their own experiences but also through the experiences of others. After seeing a person injured by fire one learns to keep away from it. Sharing and interchanging experiences, ideas, and knowledge, come out as the familiar terms: "teaching" and "learning". Teaching in the old days was in the tutorial form. Great scholars such as Socrates and Confucius had many disciples. As the world population increased and a far greater number of people wanted to

learn, education gradually evolved into the present day's school system. As the demand for education increases, the deficiency of schools and teachers becomes more obvious. This deficiency cannot be met merely by building more schools and training more teachers. Education must be more efficient.

The invention of movable type has enabled knowledge to be widely extended through the use of books. Audio-visual aids: film projectors, television sets, phonographs and tape recorders are also helpful in the modern schools and colleges. All of these serve one function for the teacher: to present material to the students. That is far from being enough for efficient teaching, however. The "productive interchange" between teacher and students in the small classroom, or in the tutorial system are completely lost. The student becomes more and more a mere passive receiver of information while information comes to him regardless of where he stands and how well he is prepared to receive the new information. The student starts from page 1 of the book or reel number 1 of the movie films. Occasionally, he encounters something he has already been quite familiar with; sometimes he encounters great difficulties. To repeat helplessly again and again something he has had difficulties with, or to spend his time on something he already knows is a waste of time. Ramo<sup>21</sup> has indicated that our technological growth is far in advance of our educational system; one method of integrating the two would be by the use of the teaching machine in a regular classroom. The teaching machine is, essentially, a device which increases the control over the learning activities

of the students. It differs from the traditional teaching devices (books, etc.) in that it plays an active role just as a teacher does during a student's course of learning. It preserves the technological advantages of printing and other devices which allow economical dissemination of knowledge, and at the same time reinstates some of the advantages of the face-to-face teaching method.

## CHAPTER 2

### CONVENTIONAL TEACHING MACHINES

Pressey<sup>17</sup> was the first person who suggested the idea of using the teaching machine. In his early article (1926), Pressey apparently foresaw his machine primarily as a teaching aid for drill and informational material. He designed a multiple-choice machine which is probably also known as the "Drum Tutor".<sup>17,18,19</sup> However, his idea was not widely recognized until as late as 1958 when statistics<sup>8</sup> have shown to be the beginning of the golden age of teaching machine research. Dozens of teaching machines have been designed. Although, technically, they are quite different, all of them implement, more or less, the same principles or methods which follow the theory of learning. This chapter is a complete survey of these ideas and devices.

#### Reinforcement

The learning theory which is strongly emphasized by Skinner<sup>23</sup> is the principle of reinforcement. When an organism behaves, it acts upon the environment. This changes the environment in some way. In turn, changes brought about in the environment feed back to the organism, and as a



consequence, affect its future behavior. When environment feedback strengthens the behavior which has brought it about, it is said that the organism has been "reinforced" or rewarded. Behavior is learned only when it is reinforced, and immediate reinforcement has a much stronger effect than delayed reinforcement does. Perin<sup>14</sup> studied this and indicated that a delay between a response and its reinforcement of a few seconds will greatly reduce the effectiveness of reinforcement. Moreover, reinforcement must be made precisely contingent upon performance of the behavior that is being taught, so that desired behavior is strengthened and unwanted behavior is discouraged. Lastly, reinforcement must be repeated for a sufficient number of times so that its effect becomes prominent.

In traditional classroom teaching, the conditions of reinforcement are very poor. A major reason is that it is physically impossible to have a teacher take care of more than thirty or forty students at the same time. Porter<sup>16</sup> has estimated that in a spelling class of thirty in an elementary school, the teacher would have to be capable of reinforcing his class about 225 times per minute in order to do an efficient teaching job. This is obviously a super-human deed. Also, whenever the teacher gave an examination, there would always be a delay of many hours or days between the examination and the return of the corrected examination paper because of the limitation of the teacher's time and effort. The effect of reinforcement then becomes greatly reduced. It is commonly observed that students are most eager

to learn of their performance in an examination right after it is over. They begin to care less and less about the examination as time elapses, because all other outside distractions divert them from it. It is no wonder that a corrected examination paper means nothing to a student if it is returned three or four weeks after the examination is over.

Without any exceptions, all the teaching machines designed emphasize the effect of immediate reinforcement. The student is informed of the correct answer to a problem (or whether his answer to that problem is correct or not) right after he finishes that problem. By pushing a button (Angell's Automatic Rater<sup>1</sup>, Briggs' Subject Matter Trainer<sup>2,3</sup>), or moving a lever (Skinner's Disc Machine<sup>24</sup>) or by pulling out a paper tab (Bryan and Rigney's "Pull Tab" device<sup>4</sup>), or looking up a certain page of the book (Crowder's Scrambled Book<sup>6</sup>), the student can obtain the "knowledge of result" with almost no time delay. In this respect, the Scrambled Book is not as desirable as the others, because looking up a certain page in the book is more or less a psychological burden in comparison with simple mechanical motions. One of Pressey's devices gives out tangible reward such as a candy life-saver. This is hardly necessary, since success, approval, and being correct have, in themselves, a strong reinforcing effect.

Immediate Feedback

Howard<sup>11</sup> points out that teaching is a two-way controlled feedback process. In one way, the teacher should be well informed about (1) the student's present knowledge of the subject, and (2) the student's learning characteristics, so that the teacher can have suitable adjustment in his teaching, according to the student's individual measure of these two factors. On the other hand, the student should be well informed about his performance and his standing, so that he can discern any errors before committing them too deeply, and be constantly aware of his progress through immediate confirmation of the correctness of his response, thus saving him from wasteful overstudy. The amount of feedback around the learner is almost zero when books or other audio-visual aids are used. The books invariably present the same kind of information to every student, while the student tries, very often vaguely, to accept the information without knowing where he stands. The situation in classroom teaching is not as bad, but it is far from being good enough. The teacher can always collect a small amount of information about his students in the class meeting, but the precise information only comes from examinations which are given probably once a month. The students are also in the same situation of unawareness. Sometimes one is not aware of what is wrong with his study until he is told that he has flunked the course.

The two-way feedback between the teaching machine and

the student is precise and immediate. Therefore, there is a constant interchange between the machine and the student. Sustained activities induced by the machine keep the student constantly alert and busy. As mentioned above, the student knows the result of his last step before he goes on for the next one; this enables him to do any correction and adjustment which is necessary. Pressey's Multiple-Choice Machine keeps an error count. Skinner's Disc Machine remembers any error the student makes and requires the student to repeat all the problem he does wrong.

#### Adaptability to Individuals

In the classroom, teaching goes at the rate which the teacher considers to be adequate to most of the students in the class. Individual differences are thus greatly neglected. Those who could go faster are penalized. They waste their time in the class, become bored, and might even lose their interest in that subject. Those who should go slower are poorly taught and unnecessarily punished by criticism and failure. Not only will they not be able to learn the subject well but they might also lose their confidence in their own abilities. The teaching machine permits every student to progress at the rate that suits him best without hampering others or being hampered by others. Brighter students can learn faster and learn more subjects. As Skinner<sup>24</sup> indicated, the grading system may then be somewhat different too. Since the teaching machine assures the mastery of every stage of

instructions during the course of learning, every student will learn the subject as well by the time he gets through. Therefore, every student can get a grade A which is an indication that his performance in that subject has been satisfactory. This is a more encouraging grading system than the currently existing one. No one is discouraged by getting C's or D's while the better students are encouraged by getting more A's than the average students.

Also the teaching machine enables the student to learn at any time he prefers. The student does not have to worry about falling behind the class because of absence. He can catch up with no trouble at all. Home study for the physically handicapped student is possible. Moreover, in industry or military training, it becomes more convenient to have individual training than to schedule the trainees in groups.

As discussed above, when the information concerning a student's present knowledge and learning characteristics feed back to the teaching machine, the machine should be able to make an internal adjustment of the teaching program so that the program will suit the individual best. The techniques often used are : repeat, skip, and branch. Skinner's Disc Machine requires the student to repeat every problem he does wrong. Crowder's Scrambled Book provides branching according to the student's answer to the previous problem. Students making incorrect responses are given remedial help or are required to repeat a certain section of the main program. These kinds of adjustment are, of course, very primitive and simple. Those at a more sophisticated level will be discussed in the next chapter.

### Different Modes of Operation

Furthermore, for higher adaptability and versatility, teaching machines are designed with different modes of operation which enable the machine to do a more efficient job under different circumstances and different training requirements. For example, one mode of operation would give the student only a brief survey of the subject matter while another mode of operation would require the complete mastery of the subject. Briggs' Subject Master Trainer<sup>3</sup> can be operated in six modes as follows:

The coaching mode: In this mode, the student merely reads the item, presses a special button which shows a green light next to the correct answer. The student then presses the correct answer button and he is permitted to go on to the next question. Notice that there is no quizzing involved.

The single error permitted mode: In this mode, the student is allowed to make only one choice. If his response is correct, a green light will be shown. However, if the response is wrong, a buzzer sounds and then a green light is shown glowing next to the correct answer.

The practice mode: In this mode, the student can make any number of errors before the correct response.

The single try mode: Similar to the single error permitted mode, but the student is not told the correct answer when he makes a wrong response. He knows only that his response is wrong.

The paced-practice mode: In this mode, an electrical

timer is connected to control the exposure time of the stimulus item. Then the machine can be operated under any of the above-described modes.

The test mode: In this mode, the machine is merely an automatic quizzer. The student makes one attempt to answer back each item, but no feedback or knowledge of result is given.

Irion and Briggs<sup>12</sup> have designed another device which has four modes of operation.

### Timing

Most of the teaching devices assume a passive position in permitting the student to take as much time as he wants in making a response. In one sense, this is fine because, in this case, the student is never embarrassed as he would be by a human teacher, urging him to go faster. However, in some programs where speed is a measure of success of the training, such as trouble shooting problems in electronic devices and control problems in system operations (Lumsdaine<sup>9</sup> has developed programs for SAGE operations), timing should be considered. Students should not only be taught to do things correctly but also be trained to acquire speed. Also, in other cases, even though speed is not of such importance, a student fumbling too long with a simple problem indicates that he might need some help. It is unfair simply to ignore this indication and wait until he painfully makes a wrong response. It should be mentioned at the same time that books

or other teaching aids functioning only to present information are unable to give any consideration to this respect at all.

Bryan's<sup>5</sup> Optimal Sequence Trainer is a device designed to give knowledge of results and guidance during long, multi-loop electronic trouble shooting problems. The student is required to trace the optimal steps to be taken. If he hesitates more than 7 minutes between steps, the device tells him the correct step to take next.

### Input-Output

Most of the teaching machines ask multiple-choice type problems. The student picks out the right answer from 4 or 5 alternatives. A few exceptions are Skinner's Disc Machine, Porter's Simple Write-In Machine, and Wycoff's Film Tutor. In Skinner's machine, a sentence is presented with one or more words omitted. The student must read the sentence and fill in the missing words correctly. He records his response on a strip of paper and compares it with the correct answer later on. Porter's machine works almost the same as the Skinner machine does. In Wycoff's machine questions are shown on a screen by means of an 8 mm. projector. The student responds by constructing his own answer on a special set of five keys. Various combinations of the keys make up the alphabet and 10 digits.

As to the question: what type of problem should the machine ask, there are different points of view. Skinner strongly favors the problems which require the student to



compose his response rather than to select it from a set of alternatives such as in a multiple choice self-rater. He believes that learning requires the student to recall and not just to recognize, since behavior is learned only when it is emitted. This is certainly true in many cases when the student does not know the material well enough to compose the correct answer of his own, yet is still able to pick out the answer from a set of alternatives among which, he knows definitely, there is a correct one. Another reason is that multiple choice material must contain plausible wrong responses which will tend to strengthen the unwanted forms, or in other words, to incur negative reinforcement. Pressey thinks that multiple-choice type problems are helpful in tracing the student's trouble. When a student makes an incorrect response, his misconceptions and difficulties are exposed. Proper remedial training can be planned by use of this knowledge. Crowder's Scrambled Book is a device which carries out this idea. The teaching program branches out according to the student's responses.

There is another reason which makes many authors prefer the multiple choice type problems to the free-response type problem. Multiple choice type problems are much more easily handled. Since there is only a finite number of alternatives, the student's responses are in a simple format and can easily be put into the machine and checked. In quite a number of devices, the student responds by pushing buttons (Angell's Automatic Rater, Brigg's Card Sort Device, Pressey's Multiple-Choice Machine) or by pulling out paper tabs (Management

Research Associates' Multi-Purpose Self Trainer<sup>\*</sup>, Bryan's Pull Tab<sup>4</sup>) or by punching holes on a paper (Peterson's Envelope and Cardboard<sup>15</sup>, Pressey's Punch Board<sup>20</sup>). An interesting and promising device worth mentioning is Peterson's Chemo Card<sup>15</sup> which is a specially prepared answer sheet for multiple-choice questions. The student responds by marking a spot with a special ink. If it is the wrong choice, the spot turns red; if it is the correct choice, the spot turns black.

Skinner's Disc Machine and Porter's Simple Write-In Machine require the student to compare his answer with the correct answer revealed by the machine and to tell the machine whether he has done the problem right. This is not quite desirable, because ambiguity concerning the correctness of the student's answer might arise if a problem has several possibly equivalent answers. Also, cheating might occur, especially with small children.

Output from the machines is mostly in printed form, with some output provided by lights and buzzers.

### Method of Operation

The teaching machine proceeds in gradual, finely-divided steps. It is a bad practice in teaching to let the student make errors constantly. When the material is so difficult that a student makes many errors, he might stop working, become irritated, and lost his interest eventually. Gradual progression not only helps the student to be correct as often

---

\*

Management Research Associates, 317 Sauk Trail, Park Forest, Illinois.

as possible, but is also the fastest way to develop a complex repertoire. Of course, over-finely divided steps are not advisable, because in this way progress will slow down and the student's time and effort will be wasted.

The teaching machine always insists on a correct answer before going on. The next problem is not revealed until the last one is correctly done. This leads the student to progress step by step, allows him to have only the material he is ready for, and helps him to form a good and effective habit of learning. When reading a book, many students have the habit of skipping over those parts with which they have difficulties and going on for the familiar parts. This is a very bad habit and it affords a great waste of time and effort. Skinner's Disc Machine allows the student to go on when he makes an error and asks the student to repeat the problems he has done wrong after he goes through the disc the first time. This is due to the limitation of the mechanical structure of his machine. Teaching Machines Inc.'s Omnideck\* which is simply a set of cards containing a question on one side and the answer on the other side will not be able to prevent the student from glancing ahead and is not desirable in this respect.

---

\* Teaching Machines Incorporated, Box 28, Albuquerque, New Mexico.

CHAPTER 3THE COMPUTER AS A TEACHING MACHINE

All the teaching machines are built with a common purpose: to avail the student of the benefits and conveniences of having a good, private tutor. To a certain extent this purpose is achieved in those conventional teaching machines. However, they are still incomparable to a human teacher in many respects. Of course, we cannot and will not expect a simply constructed machine (as most of the teaching devices are) to imitate many of the complicated human's behavior: i.e., judging, adapting, and adjusting, which affects a great deal the success of a teaching program. We are looking for a device which will do a more efficient and competent job. A high speed computer is a good choice. By using a computer as a teaching device, we can expect that it will not only do as satisfactory a job as that of a human teacher, but will exceed him in many instances.

Keislar<sup>13</sup> calls that type of teaching machine which uses a computer with a large storage mechanism "IDEAL". He says that this is one of the two theoretical extremes of teaching machines. (He calls that type of machine which is a "successive approximations model" the other extreme.)

Skinner is the one who develops and advocates this "successive approximations model". Most of the authors also tend to do their research work at this extreme. Rath's<sup>22</sup> work is the only known research project which tries to use a computer (IBM 650 computer) as a teaching device. Some features of using the computer as a teaching machine are investigated as follows:

### Speed

The electronic computer is a very high speed device, operating at a rate of microseconds or milliseconds. Let us take an IBM 704 computer as an example; it can execute instructions at the rate of about 40,000 per second on most problems. This tremendously high speed permits the computer, when used as a teaching machine, to do a job which is beyond the reach of any other teaching device or of any human teacher.

When a student does a problem and related his answer to the computer, it will take only milliseconds for the computer to check this answer and to inform the student of his accuracy. In a typical program prepared for an IBM 704 computer, it takes less than 20 instructions to check the answer, count the errors, and get ready to print out a statement telling the student of the result. These will take less than 2 milliseconds altogether (excluding the time of printing out the statement on the flexowriter). Therefore, what the student actually feels is that the machine responds to his answer with no hesitation at all. In the conventional teaching

machine, it takes at least several seconds to move the lever or to pull the paper tabs. Even a human teacher's thinking is slower than the computer. As discussed above, the principle of reinforcement requires immediate reinforcement for efficient learning. A computer really does the fastest possible job.

If the machine is teaching a mathematical subject, it is not even necessary to have the answers of all the problems calculated in advance and stored in the memory space. We can program the computer to print out a problem and set it to function simultaneously as the student works. The computer will always get the correct answer before the student does, and can "wait" there to check the student's answer.

Moreover, when the student records his answer, digit by digit or letter by letter, the computer runs so fast that it can be programmed to check the accuracy of every digit (or letter) before the next one is written. If the computer finds a digit (or letter) is wrong, it will inform the student immediately before he writes down the next digit or letter and prevents him from going any further. This not only saves the time wasted in completing an answer which has already been checked out and proved partially wrong, but also tells the student exactly which digit (or letter) is incorrect.

### Storage

Conventional teaching machines might contain 30-40 or up to 100 problems. For instance, Skinner's Disc Machine has about 30 frames on every disc. This is very limited. A

computer provides a huge amount of memory space. The IBM 704 computer, as an example, contains 32,000 core storage registers (each register can store 6 numerals or alphabets), 8 magnetic drums (8,192 words per drum unit), 10 magnetic tapes (96,000 words per tape unit). With these memory spaces we can put as much information as we like in the teaching machine. We can have as many problems as necessary, so that the teaching proceeds in a gradual, finely divided gradation. We can have different types of problems, so that each individual will have problems which suit him best. Extra explanations of certain difficult topics and different demonstrations looking at a topic from different points of view can also be available, so that they can be presented to the student whenever needed. References or related information can be listed similar to a library catalogue and will be given out differently to different individuals. In short, a computer is similar to, or perhaps even better than, a very learned teacher with all the materials about the subject at his disposal, and thus it could, of course, do a magnificent teaching job.

Moreover, the computer can keep a detailed account of the student: not only will it record the number of errors made and the total grade, but past performances in every lesson or every section, trouble areas, characteristics of the trouble, etc. All these facts will be very useful in analyzing the student's present knowledge and learning characteristics which will lead to a suitable adjustment of the teaching program.

## Logical Operations

The computer can perform logical operations which no other teaching devices can. This ability greatly enhances the computer's potential as an ultimate substitution of the human teacher. As mentioned above, a computer can keep a detailed record of the student's performance at every stage. The computer can analyze the student's present knowledge and learning characteristic based on some formulas and criteria, like counting the errors, grouping the mistakes the student has made, finding the weighted sum of certain performance indexes, etc. This information is useful in adjusting the teaching program. Moreover, the computer can relate information to the student telling in which area he is weak or is proficient and what are the inclinations of his learning characteristics ("sharp but careless", "good grasp of abstract ideas", etc.). Thus some valuable personal adjustment can be made.

The computer can then modify the teaching program as teaching proceeds. Keislar<sup>13</sup> calls such a device an ideal ultimate machine in which each step would be dependent upon everything that has gone before. Techniques to modify the teaching program such as skip, branch, and repeat can easily be employed in a computer. In a private discussion, Ronald A. Howard has suggested applying the method of dynamic programming to help modify the teaching program. At a branching point, the computer should have a list of all the possible present states (states of knowledge) that the student might be in, and a list of all the possible alternatives (alternatives of



teaching program) in every state. The computer will analyze the student's past performances, decide the student's present state, generate the transition probabilities and rewards of every alternative in every state (which are functions of past performances) and find out what the best teaching program to be followed is. (For details, see "Studies in Discrete Dynamic Programming", Doctoral Thesis of Ronald A. Howard, May 1958, M.I.T.) When a teacher wants to modify the teaching program, he makes his decision only through very rough estimation. However, in a computer, every step is precise and accurate.

Not only is the teaching program determined by the student's performance, but also the goal as well. This means that the goal of the teaching machine is not decided at the beginning of the program. However it might be partly pre-determined and partly unfixed. For example: if a student performed well he would be taught some more advanced topics, while a poor student would be taught only the basic materials. A student who shows strong interest in practical works will be taught the more practical aspect of the subject and a student who enjoys theoretical materials more will be instructed with greater stress in this area. A student will learn a subject in the most effective way and learn the materials that conform with his interest and ability.

Like Briggs' Subject Master Trainer, a computer can also be programmed to have different modes of operation. There would be several parallel programs, each of which would teach the subject in a different mode. A different mode of operation would be selected by entering a corresponding selecting message.

## Input-Output

The student and the computer can communicate through some suitably connected terminal equipment. A flexowriter would be quite adequate for this purpose. The computer can print out problems or any information on the flexowriter, and the student can compose his answers in both literal and numerical forms. Very few computers have a built-in flexowriter as an input-output unit. In that case, it is necessary to prepare a "utility program" which connects the flexowriter to the computer. (A Flexowriter Monitor Interpreter System is being developed by the M.I.T. Computation Center.)

By use of the flexowriter, the computer can do a much more sophisticated teaching job than the conventional teaching devices. Since the computer can recognize any text written down by the student, it can check the student's answer whether it is a multiple choice type problem (student writes down his choice among 1,2,3,4,5.), or a self-constructed type problem (student composes his own answer which might either be English text or numbers). For those authors who prefer self-constructed type problems, there could be no necessity to ask the student to check his answer by himself (as the Skinner's Disc Machine). This will have a strong effect in easing the student's psychological burden during the course of learning.

Besides the flexowriter, the oscilloscope can be used to display graphs, diagrams, or maps which would help the

student to understand the subject. A number of computers, for example: IBM 704 and TX-O, have an oscilloscope as an output unit. It would also be possible to connect to the computer a tape recorder or a phonograph which, when triggered by an output signal from the computer would give the student instructions or explanations in vocal form. These would add more incentive to the teaching and would make the teaching more efficient.

The problem of providing the student with some other input equipment other than the flexowriter is a more complicated one. In order that the student may put his answer in graphical form, the problem of pattern recognition must be solved. The computer must be made to interpret the verbal statements made by the student so that the student may answer orally. However, once these problems are solved, the computer will undoubtedly be a perfect teaching device.

### Timing

Like Bryan's Optical Sequence Trainer, a computer can have a timing device (in the Flexowriter Monitor Interpreter System developed by the M.I.T. Computation Center there is an alarm clock connected to the computer). The computer can keep the time after a problem is presented. If the student does not respond within a certain period, it is an indication that he has encountered some trouble and the computer would give him the necessary help. Moreover, the computer can find out how long the student takes to do a problem. This

is also valuable information concerning his present state of knowledge. Since there is obviously a difference, if one student can get the right answer in 1 or 2 minutes while another student takes 10 or 20 minutes to do it correctly.

### Method of Operation

The computer proceeds step by step. In no way can the student violate the pace of proceedings. The student must get the correct answer to one problem before he can go on for the next. He can never glance ahead, because everything is "deep" in the core memory of the computer. He cannot cheat because the computer checks the student's answer itself and would never compromise with a wrong answer.

### Cost

The main factor which prevents most of the authors from attempting to use the computer as a teaching device is its prohibitively high price. One hour of machine time on an IBM 704 Computer would cost about \$500 to \$600. If we take this amount as the expense of teaching one student per hour, the idea of using the computer as a teaching device appears to be completely unrealistic. However, when we examine the situation more closely, we find that only milliseconds are spent for the computer to check the student's answer and to make decisions about any modifications of the teaching program. Most of the time is consumed by the student in thinking, composing his answer and recording on the flexowriter. When

the student is thinking and composing his answer, the machine stands idle and all the machine time is completely wasted. Also, when the student writes his answer on the flexowriter, there is always some time lapse between the interval of striking two keys in succession. Although this is only in the order of a fraction of a second, it is quite a long time for a high speed computer. It has been estimated that when one character is recorded on the flexowriter an IBM 704 computer would take only a few hundred microseconds to receive the information and store it in the memory space. The difference between a fraction of a second and a few hundred microseconds is the machine time wasted in writing one single character. This accumulates to a considerable amount of machine time when a sentence or a paragraph is written. Therefore, although a student would spend 1 or 2 hours to learn a subject, the machine time he uses might just total to several seconds. If a time sharing scheme is worked out, the cost charged on every student will then be very little. In order that more than one student can be taught at the same time, the computer should have as many input-output channels as the number of students so that every student can communicate with the computer through his own terminal equipment without interfering one another. In a rough estimation, it is said that since, in general, less than one second of machine time on an IBM 704 computer is needed to take care of a student for one hour, (the 704 computer can carry out about 40,000 instructions every second), a 704 computer can teach up to 2,000-3,000 students simultaneously.

This cuts the cost down to less than 50¢ on a per student per hour base . Furthermore, it is believed that if a special purpose computer can be developed to give special considerations on the input-output- channels, time sharing would be carried out in a more convenient and more efficient way. When the Flexowriter Monitor Interpreted System is used, the computer can teach a student and, at the same time, perform some other unrelated computation work when the student is thinking and writing down his answer. These factors would greatly cut down the cost on a per person base, and thus make the idea of using the computer as a teaching machine practicable.

CHAPTER 4A PROGRAM TEACHING MATRIX ALGEBRA

A program to use an IBM 704 computer to teach a course in matrix algebra is prepared. The main purpose is to investigate how a computer can do the teaching works which are usually done by a human teacher. Therefore, in this program, teaching is conducted in a way which is very close to the present day classroom teaching (Another program which follows the learning theory and works very similar to the conventional teaching machines is discussed in the next chapter.) The student comes to the computer, gets his reading and problem assignments, finishes it before he comes to the class again the next time, and takes an examination in every lesson. This is just like attending a huge class where the feedback loop around the learner is rather big. A comparison with the program in the next chapter will reveal how different teaching principles and teaching methods are applied in different cases.

Contents

The course uses the first chapter on matrices, determinants, and linear equations of "Methods of Applied Mathematics" by

Hildebrand as the text book. Several sections on numerical methods are skipped (following the schedule of the course 18.15 which is taught at M.I.T.). The material is divided into 10 lessons.\* This requires the student to come to the class eleven times to finish the course. The problem set in the book is adopted. A supplementary problem set is suggested to give the students more practice when it is necessary. Every lesson has two sets of quiz problems; each set contains 10 multiple-choice type problems. Multiple-choice type problems are used because the flexowriter system has not yet been developed to allow the student to write down his answers in the form of a matrix or a polynomial, etc.

#### Checking and Grading

The computer checks every quiz problem the student does and grades it. The student chooses an answer from five alternatives, one of which is: "I DO NOT KNOW EXACTLY WHICH IS THE CORRECT ANSWER". In this way, instead of picking out an answer blindly (a very bad practice) the student can indicate that he is unable to decide. The grading system is: 5,3,1,0,0 for every problem, e.g., there is a correct answer (5 points), a partially-correct answer (3 points), an answer which shows some understanding of the materials (1 point), an incorrect answer (0 point), and an undecided answer (0 point). After checking a problem, the computer prints out "THE CORRECT ANSWER TO PROBLEM NO. (2) is (3). YOUR GRADE IS (1) OUT OF

---

\* See Appendix A.



5 ". When a quiz problem set is finished, the computer prints out "YOUR TOTAL GRADE IN THIS EXAMINATION IS (32) OUT OF 50 ". When the complete course is finished, the computer prints out "YOUR AVERAGE GRADE IN THIS COURSE IS (38) OUT OF 50 ".

In every examination, 50 is the highest grade and 30 is the passing grade. If a student scores below 30, he is asked to do some review problems and take the examination again (there are two sets of examination problems in every lesson). If he fails twice, he will continue with a low score in his record.

### Troubles

If a student has any trouble with the course, the machine will look into his trouble and give him a suitable problem assignment, so that he can have more practice in that specific part before he goes on. When the student comes to the class, the machine will print out first "PLEASE ANSWER BY YES OR NO WHETHER YOU HAVE FINISHED YOUR LAST ASSIGNMENT FOR LESSON (FOUR)". If the student answers in the affirmative he can go on to take the examination for that lesson, but if his reply is negative, all the possible troubles he might have had in that lesson\* will be listed by the machine and the student will be required to indicate the one applicable to his situation. Here again, the student can answer the question in a multiple-choice fashion only. This is because the machine would have

---

\* See Appendix A.

to be able to understand any written statement in English if the student were permitted to describe his difficulties in his own words. This is still an unsolved problem for many linguists who are interested in having a machine which can understand freely composed English text. In order that the student's trouble can be evaluated analytically, a weighting point\* is given to every reason of trouble in the list according to the degree of difficulty of that specific part of the materials. Among all the possible troubles the student might have in a lesson, there is always one which reads: "I DO NOT HAVE ENOUGH TIME TO FINISH MY ASSIGNMENT". If the student chooses this response it will be noted by the machine which will print out one of the following appropriate instructions.

(1) THIS IS THE FIRST TIME THAT YOU ARE TOO BUSY TO FINISH YOUR ASSIGNMENT. YOU ARE EXCUSED AND MAY FINISH THE ASSIGNMENT BEFORE THE NEXT CLASS.

(2) THIS IS THE SECOND TIME THAT YOU ARE TOO BUSY TO FINISH YOUR ASSIGNMENT. YOU ARE EXCUSED AND MAY FINISH THE ASSIGNMENT BEFORE THE NEXT CLASS.

(3) THIS IS THE THIRD TIME THAT YOU ARE TOO BUSY TO FINISH YOUR ASSIGNMENT. IT IS ADVISED THAT YOU SPEND SOME MORE TIME ON YOUR WORK. FINISH YOUR LAST ASSIGNMENT BEFORE YOU COME TO THE NEXT CLASS.

(4) THIS IS A WARNING THAT YOU HAVE FAILED TO FINISH YOUR ASSIGNMENT MORE THAN THREE TIMES. BE SURE TO HAVE YOUR ASSIGNMENT FINISHED BEFORE YOU COME TO CLASS HEREAFTER.

---

\* See Appendix A.

## Records

The computer keeps a complete record of the student which includes: (1) his grade in every examination he has taken, (2) his trouble (counted in weighting point) in every previous lesson, (3) his trouble (counted in weighting point) in the present lesson, (4) what part (or parts) of the present lesson he has had trouble with, (5) whether he has taken an examination (and failed, of course) in the present lesson, and (6) his total count of being unable to finish his assignment just because he had been too busy.

## Flow Graph

The flow graph of the program is shown on the following pages. Several features about the program:

Every student has the same reading assignment, but the problem assignment may vary for different individuals. Every lesson should include:

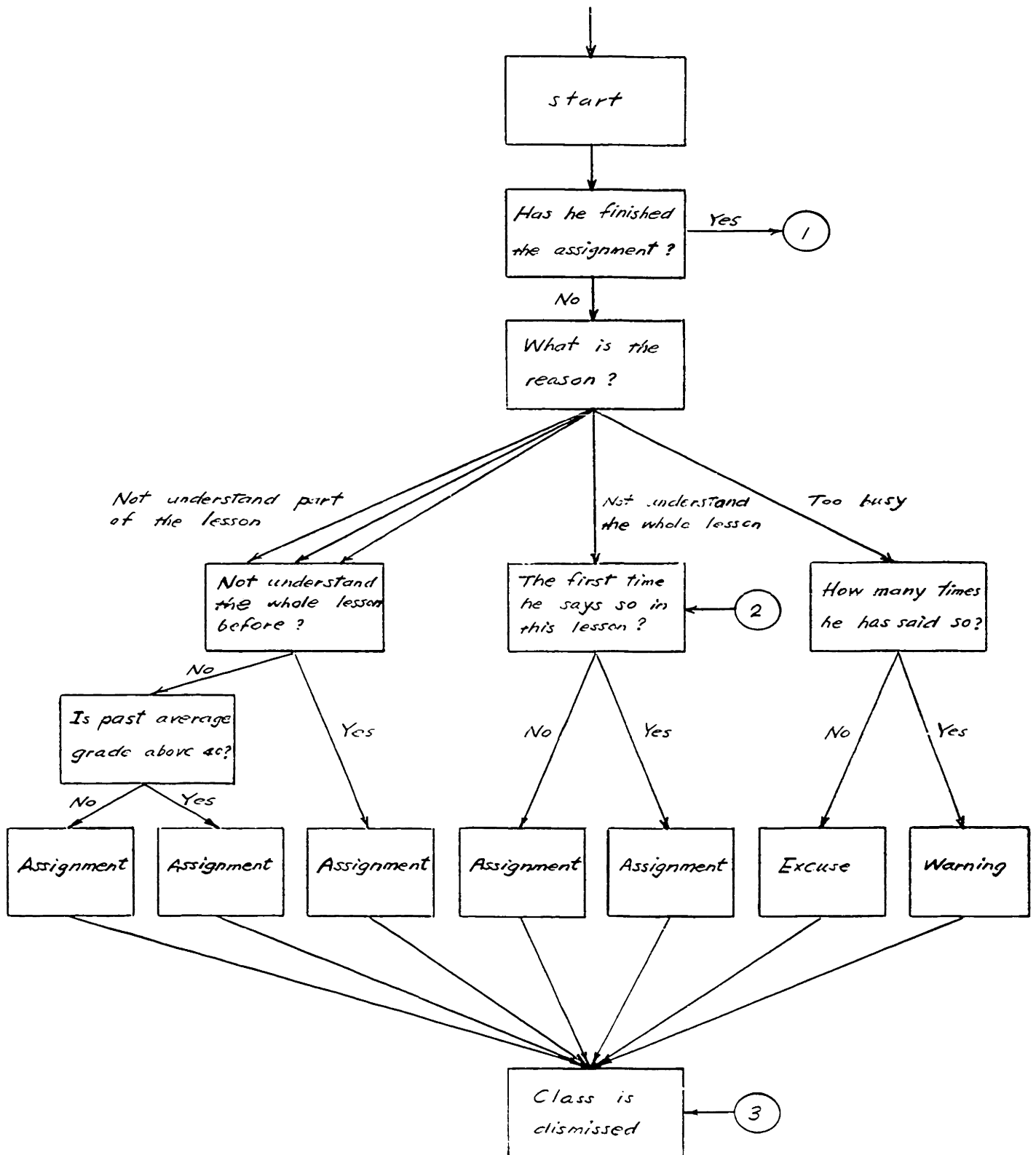
For the students who take and pass the examination:

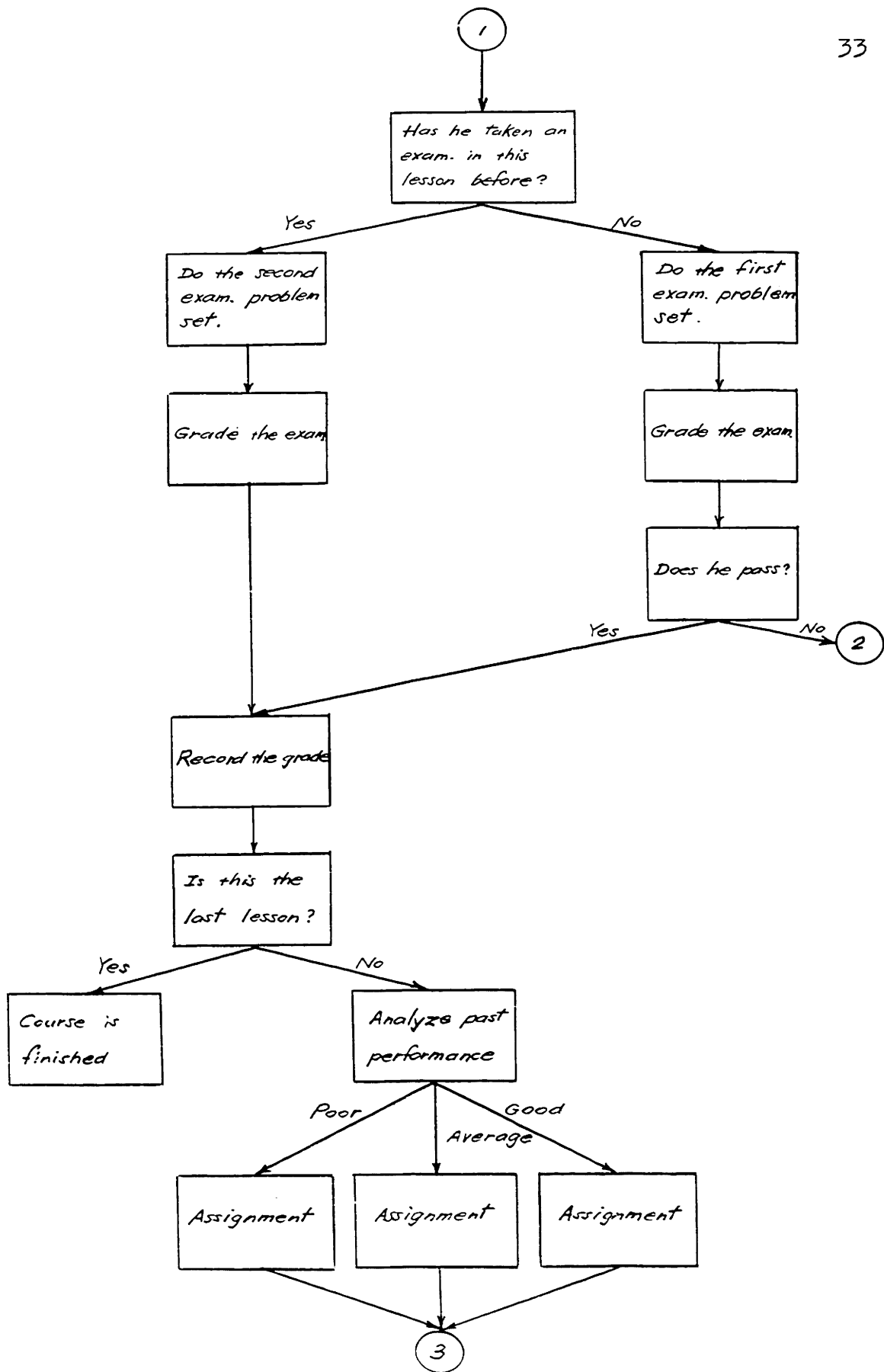
1. A problem assignment for good students
2. A problem assignment for average students
3. A problem assignment for poor students

For those students who do not finish their last assignment or who take the examination and fail the first time:

4. A problem assignment for students who indicate, for the first time, that they do not understand the lesson completely

5. A problem assignment for students who indicate, for





more than one time, that they do not understand the lesson completely

6. A problem assignment for students who have difficulty with part 1 of the present lesson, but whose past performances are good (average grade above 40)

7. A problem assignment for students who have difficulty with part 1 of the present lesson, but whose past performances are poor (average grade below 40)

8. A problem assignment for students who have trouble with part 1 of the present lesson, but have had trouble with the whole lesson before.

(Repeat 6,7,8 for part 2,3.... of the lesson)

When a student takes an examination and passes it, his assignment for the next lesson is given out based on his past performance. His past performance is analyzed by the formula:

$$x = y \text{ (weighted average grade in every past lesson) } - Z \text{ (weighted average trouble in every past lesson)}$$

If  $x \geq 38$ , he is a good student. If  $38 > x \geq 30$ , he is an average student. If  $x < 30$ , he is a poor student. The weighted average grade in every past lesson is calculated by weighting the grades in the recent lessons heavier and weighting the early lessons lighter:

$$Y = \left[ \begin{aligned} &2.0 \times \text{(grade in the present lesson)} + \\ &1.8 \times \text{(grade in the previous lesson)} + \\ &1.5 \times \text{(grade in the second lesson preceding the present} \\ &\quad \text{lesson)} + \\ &1.2 \times \text{(grade in the third and fourth lesson preceding} \\ &\quad \text{the present lesson)} + \\ &1.0 \times \text{(grade in all the earlier lessons)} \end{aligned} \right] \\ \div (2.0 + 1.8 + 1.5 + 1.2 + 1.2 + 1.0 + 1.0 \dots)$$

Similarly

$$\begin{aligned}
 Z = & \left[ 2.0 \times (\text{trouble in the present and previous lesson}) + \right. \\
 & 1.5 \times (\text{trouble in the second and third lessons preceding} + \\
 & \quad \text{the present lesson}) \\
 & 1.0 \times (\text{trouble in the fourth and fifth lessons preceding} + \\
 & \quad \text{the present lesson}) \\
 & \left. 0.5 \times (\text{trouble in all the earlier lessons}) \right] \\
 & \div (2.0 + 2.0 + 1.5 + 1.5 + 1.0 + 1.0 + 0.5 + 0.5 + \dots\dots\dots)
 \end{aligned}$$

If a student takes an examination and fails for the first time, the response of the machine will be equivalent to that situation in which the student has said "I do not understand the lesson completely" and he will be given an assignment which allows him to review the whole lesson once again.

## CHAPTER 5

### A PROGRAM TEACHING ARITHMETIC

Following is the description of a program in which an IBM 704 computer is used as a device to teach a course in addition. In this program, the feedback loop around the learner is small, the effect of immediate reinforcement is emphasized and individual differences are well taken care of.

#### Contents

The course is divided into five parts: part I introduces the idea of addition, teaching the student to add two numbers by counting one after the other. Part II teaches the student to add two numbers with more than 1 digit by adding the unit digit to the unit digit and the tens digit to the tens digit, etc. It also introduces the addition of two numbers with different numbers of digits. Part III teaches the student to add several numbers with one or more digits. Part IV introduces the idea of carry, when the sum of two numbers exceeds 10. Part V teaches the general case of addition of several numbers with several digits. In every part, practice problems are prepared which permit skip, branch, and repeat as will be



discussed below. There are six practice problems and one group of five review problems for part I, 15 practice problems and two groups of review problems, five each, for parts II, III, and IV respectively, and 15 practice problems and one group of 5 review problems for part V.\* In the whole course, a student will do a minimum of 46 problems and a maximum of 106 problems. It must be mentioned that the teaching material in this program is prepared for adult education rather than for teaching young children. Since it is not possible to try to teach a small child the whole subject on addition within a couple of hours, the gradation should be finer and more practice problems should be provided in that case.

### Checking and Grading

The computer checks every problem the student does. If he does it correctly, the machine will print out the next problem so that he can continue. If he does it wrong, the machine will print out: "YOUR ANSWER IS WRONG. TRY AGAIN", and let the student repeat until he gets the correct answer. If the student gets the problem right in the first attempt, he gets one point. If he gets it right in the second attempt, he gets zero point, minus one in the third attempt, etc. The grading system is a reasonable one. Since all the problems are fairly easy, it can be expected that an average student would do it right within one or two tries. When he does it right in the first try he gets a credit for that. When he

---

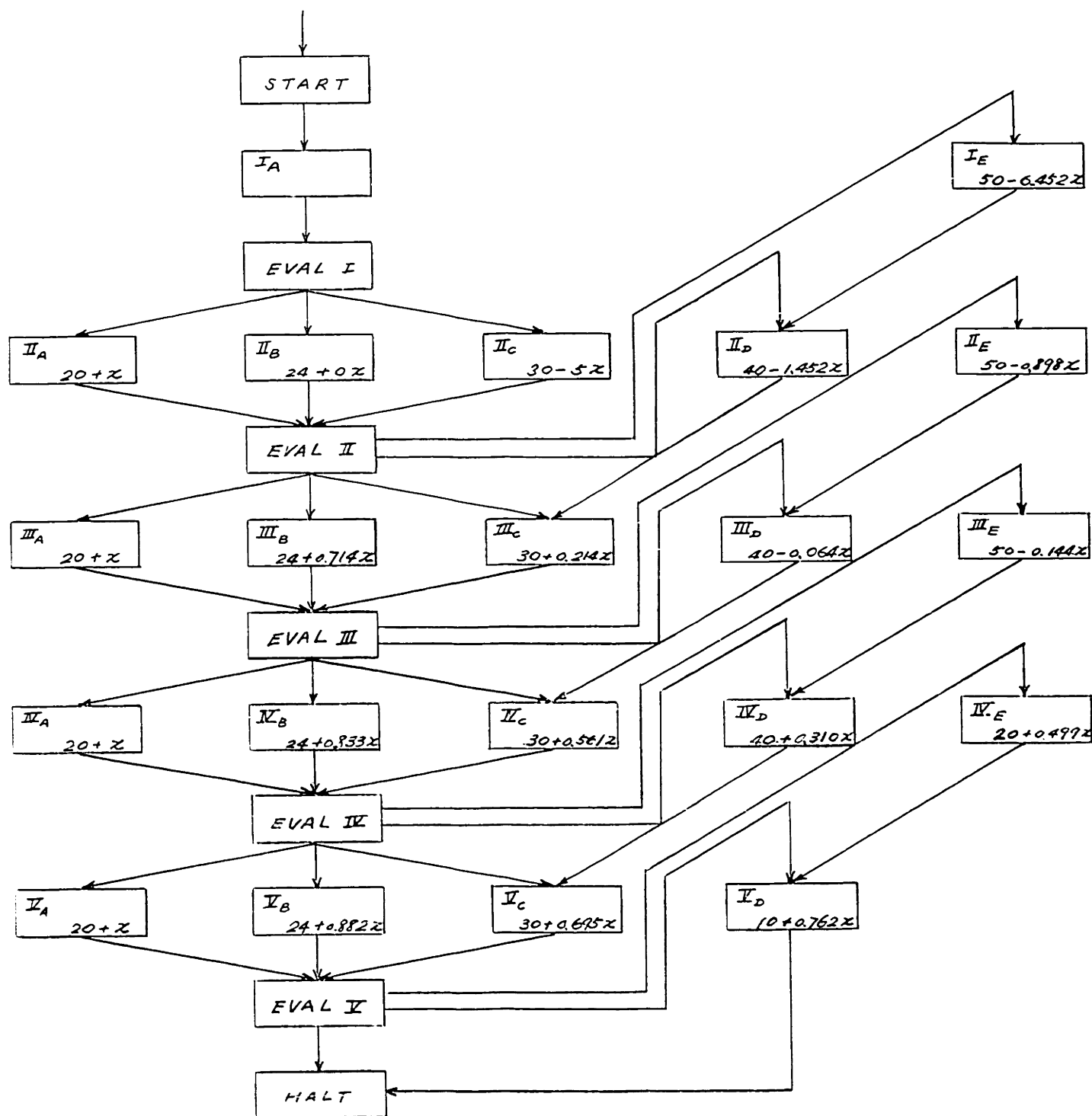
\* See Appendix B.

takes two tries to get the problem right he does an average job and gets neither credit, nor punishment. If he takes more than two tries he is below average. The grade for every problem is accumulated as the student goes on. This information regarding the student's present state of knowledge and past performance will be used in deciding how the teaching program should be modified; however, it will not be revealed to the student.

#### Modification of the Teaching Program

When a student finishes the problems in one part, his accumulated grades will be reviewed and the teaching program adjusted accordingly. The flow graph which shows how the teaching program might be changed is shown on the following page.

All students commence with the same six practice problems in part I. After that each student's performance is evaluated and put into one of three groups: A, B, C, accordingly. Group A are those with superior performances and will do ten of the fifteen practice problems in part II. Group B are those with good performances and will do twelve out of the fifteen practice problems in part II. Group C are those with average or below average performances and will do all fifteen of the practice problems in part II. This is a scheme for good students to skip part of the practice problems. After finishing the problems in part II, the student's performance is evaluated again and he is put into one of the five groups: A, B, C, D, E. If he is in group A, B, or C, he will do ten, twelve, or all



fifteen of the practice problems of part II. If he is in group D, he will do the first group of review problems (five of them) in part II, and then, regardless of how well he does the review problems, must solve all fifteen of the practice problems in part III. If he is in group E, he will do the five review problems in part I, the first group of five review problems in part II, and then all fifteen of the practice problems in part III. This is a scheme for poor students to repeat the materials which they obviously have not learned well. The cycle repeats for part II, IV, and V. When the student is finished, the machine will print out: "THIS LESSON IS FINISHED. THANK YOU", and dismisses the class. Note here the policy on modifying the teaching program is let the student do a certain number of practice problems and judge his performance; if his performance is good enough, he can go on for the next part, the number of practice problems he will have in the next part is also determined by his performance; if poor, he would be required to review the present materials and then go on; if he is very bad, it is a good idea to let him review the last part of the material and then the present material, before he goes on. Notice that there are two groups of review problems each in part II, III, and IV which allow the student to have the chance to review the material twice if he has difficulty with it.

A student's performance is evaluated by his accumulated grade. He will be put into one of the five groups which have a teaching program that will benefit him most. To determine

which group will best benefit the student, a criterion is set by assigning to each group an "intrinsic linear operator". Given the student's grade, an operator will compute the "index number" which is a measurement of the benefit the student will gain. The student will then be put into the group which gives the highest index number.

Let  $x$  denote the student's accumulated grade. The intrinsic linear operator is in the form  $\alpha + \beta x$  where  $\alpha$  and  $\beta$  are two constants.  $\alpha$  is an estimation of how much a student would be benefited by looking at how much practice that group will offer without considering past performance. For example, group A requires the student to do ten of the fifteen practice problems, group C requires the student to do all fifteen of the practice problems, and group E requires the student to review the last part before going on. If we ignore the past performances of the student, we will naturally think group E would be the best choice, group C the second choice, and group A the worst choice, since more review and more practice will definitely do a student more good. Therefore group E should have the largest  $\alpha$ , and group A the smallest.  $\alpha$  is always positive, not only because the  $\alpha$ 's are only compared with one another on a relative basis so that a positive constant can always be added to make all of them positive, but also because we believe that the student will always gain some positive benefits through practice and it is meaningless to make  $\alpha$  negative.  $\beta$  is an estimation on how much a student would be benefited by the practices offered in a group when his past performance is considered. A good student is apt to

be put into a group which requires less practice work, since this group will save his time and effort and thus benefit him more than another group which wastes his time and effort in doing something that is not necessary for him to do. Therefore, group A should have the largest  $\beta$  and group E the smallest (algebraically).  $\beta$  might be either positive or negative. A group that requires less practice work would have a positive  $\beta$ , so that a good student (whose grade is high and positive) would gain a large benefit when being put into this group, while a poor student (whose grade is low or even negative) would gain a small benefit or even incur a loss (negative benefit) because he is doing something which is not adequate for him to do. A reviewing group would have a negative  $\beta$ . A good student is not supposed to be put into this group. The better he is (the higher his grade is), the more he will lose in the waste of time and effort. A poor student is apt to be put into this group for the worse he is (the more negative his grade is), the more benefit he will draw from the review problems.

The values of  $\alpha$ 's and  $\beta$ 's of all intrinsic linear operators are labeled on the flow graph on page 39. As an example: if a student comes to EVAL II with an accumulated grade 13, that is:

$$x = 13$$

$$\text{III}_A: 20 + x = 33$$

$$\text{III}_B: 24 + 0.714x = 33.286$$

$$\text{III}_C: 30 + 0.214x = 32.786$$

$$\text{II}_D: 40 - 1.452x = 21.119$$

$$\text{I}_E: 50 - 6.452x = -33.880$$

he should be put into group III<sub>B</sub>. If the student's grade is 5, that is,

$$x = 5$$

$$\text{III}_A: 20 + x = 25$$

$$\text{III}_B: 24 + 0.714x = 27.171$$

$$\text{III}_C: 30 + 0.214x = 31.071$$

$$\text{II}_D: 40 - 1.452x = 32.738$$

$$\text{I}_E: 50 - 6.452x = 17.738$$

he should be put into group II<sub>D</sub>. When two groups give the same index number, we always go to the "higher" group, i.e., we prefer group A to group B, group B to group C, etc., in case of a tie.

Notice that the grading system, the way that the teaching program skips, branches and repeats, and the values of the intrinsic linear operators in this program are designed to fit one another. When a student performs well and gets a high grade in one part, he will be put into group A when he goes to the next part. He does fewer problems and has a smaller chance to accumulate a higher grade. Unless the student continues to have a very good performance, he will easily drop into group B or C. On the other hand, an average student might be put into group A or B on the next trial. Even when a student with a very low accumulated grade is put into group E, he still has a chance to be put into group A when he goes to the next part. Because he will do 25 problems (say, 5 in I<sub>E</sub>, 5 in II<sub>D</sub>, 15 in III<sub>C</sub>) he gets a chance to accumulate a higher grade (the highest total grade he can accumulate will be approximately the same as that one which

can be accumulated by a student in group A, because a student in group E would start with a much lower accumulated grade for his past performance). The values of the intrinsic linear operators must be suitably chosen to assure the equality of the chances of every category of students.\* Therefore, the machine keeps a watchful eye on the good students and urges them to keep up their good performance, while it encourages the poor students by letting them have the chance to get ahead. Following are some illustrative cases.

A student comes to EVAL II with an accumulated grade of 15. He is put into group A where he does only 10 problems. If he accumulates 9 points or more out of these 10 problems he will be put into group A again, when he goes to EVAL III. If he gets less than 9 but more or equal to 7 he will drop into group B. Notice that it is not very easy to keep in group A.

A student comes to EVAL II with an accumulated grade of 13. He is put in group B where he does 12 problems. When he goes to EVAL III, he will be put into group A if he gets 11 or more, in group B if he gets 9 or more. This is difficult.

A student comes to EVAL II with an accumulated grade 0. He is put in group E where he does 25 problems altogether. When he goes to EVAL III he will be put into group A if he gets 24 or more, group B if he gets 22 or more. Notice that he still has a fairly good chance.

It is true that in this program, the margin of being put into group A is very narrow. Because this is a very fundamental subject, it would be a better idea to let the student have more

---

\* See Appendix C.



practice than to let him skip unless he is really brilliant. In fact, the problems in group C require a normal amount of work and are suitable for most students. Therefore, the program intends to put most students into group C by providing adequate values of the linear operators. The point we want to make clear is that the criterion of modifying the teaching program depends strongly on the subject materials and the way the practice problems are prepared. There should be no fast and rigid rule.

## APPENDIX A

A. Reading Assignments

## Lesson 1

Section 1.1 Introduction. Section 1.2 Linear Equations and the Gauss-Jordan reduction. Follow the example on page 3 to see how the Gauss-Jordan reduction proceeds. Notice the different cases when a set of linear equations possesses no solution, or an  $(n-r)$ -fold infinity of solutions, or a unique solution. Section 1.3 Matrices. See how to find the product of two matrices. Remember that multiplication of matrices is, in general, not commutative. Try an example on how multiplication can be carried out by partitioning of the matrices. Section 1.4 Determinants and Cramer's Rule. Distinguish the definition of minor and cofactor. See how a determinant is expanded by the Laplace expansion formula.

## Lesson 2

Section 1.5 Special Matrices. See how the adjoint of a matrix is defined. Notice the result of premultiplying and postmultiplying of a matrix by a diagonal matrix. Section 1.6 the inverse matrix. Be careful about how the inverse matrix is defined. Look at equation (35) on page 14. Section 1.7 Elementary Operation. Section 1.8 Solvability of Sets of Linear Equations. Relate the result in this section to that mentioned in Section 1.1.

## Lesson 3

Section 1.9 Linear Vector Space. Section 1.10 Linear Equations and Vector Space. Notice the conditions when a set of homogeneous equations or a set of nonhomogeneous equations possesses a solution.

## Lesson 4

Section 1.11 Characteristic-Value Problems. Notice that many of the statements in this section are only true for symmetric matrix and are not necessarily true for

nonsymmetric matrix. See how the solution of a set of nonhomogeneous equations is expressed in equation (89).  
Section 1.12 Orthogonalization of Vector Sets.

#### Lesson 5

Section 1.13 Quadratic Forms. See what is a modal matrix and a normalized modal matrix. Section 1.14 a Numerical Example. Section 1.15 Equivalent Matrices and Transformations. Distinguish the definitions of orthogonal transformation, congruence transformation and similarity transformation.

#### Lesson 6

Section 1.16 Hermitian Matrices. Compare the properties of a Hermitian matrix to that of a real symmetric matrix which are mentioned in previous sections. Section 1.17 Definite Form.

#### Lesson 7

Section 1.18 Discriminants and Invariants. For a symmetric matrix to be positive definite, it requires that (i) all diagonal elements of the matrix be positive, (ii) the determinant of the matrix be positive, (iii) all the leading principle minors be positive. Section 1.19 Coordinate Transformations.

#### Lesson 8

Section 1.20 Diagonalization of Symmetric Matrices. A symmetric matrix is diagonalized by an orthogonal transformation where the transformation matrix is its modal matrix. The resultant diagonal matrix displays its characteristic values along the diagonal. Section 1.21 Multiple Characteristic Numbers.

#### Lesson 9

Section 1.22 Functions of Symmetric Matrices.

#### Lesson 10

Section 1.26 Characteristic Numbers of Nonsymmetric Matrices. Notice the orthogonal relations between characteristic vectors of a nonsymmetric matrix, which is different from that of a symmetric matrix.

B. List of Possible Reasons for Being Unable to Finish the Assignment.

Lesson 1

1. I do not have enough time to finish my assignment.
  2. I do not understand all the materials in the assignment.
  3. I do not understand the Gauss-Jordan reduction method.
  4. I do not know well about manipulating of matrices.
  5. I do not know how to apply Cramer's rule to solve simultaneous equations.
- (20,10,3, 3, 3)\*

Lesson 2

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not understand the discussion about the rank of a matrix, of an augmented matrix, etc.
4. I am not clear about how the solvability of a set of equations is related to the rank of its coefficient matrix and its augmented matrix. (20, 10, 4, 4)

Lesson 3

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not understand the section on Gram determinant.
4. I do not understand how the solvability of a nonhomogeneous set of equations is related to the solution of the transposed homogeneous set as stated on page 29. (20, 10, 4, 4)

Lesson 4

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I am not clear about how the solution of a nonhomogeneous set of equations is expressed as in equation (89).
4. I do not understand the Schmidt orthogonalization procedure. (20, 10, 4, 6)

---

\*Weighting points

## Lesson 5

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not know how to reduce a quadratic form to its canonical form. (20, 10, 6)

## Lesson 6

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not understand the procedure to reduce two quadratic forms simultaneously to canonical forms. (20, 10, 6)

## Lesson 7

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I am not clear about the conditions for a quadratic form to be positive definite.
4. I do not understand how coordinate transformation is carried out by the transformation matrix. (20, 10, 5, 5)

## Lesson 8

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not know the procedure to diagonalize a symmetric matrix.
4. I do not know how to diagonalize a symmetric matrix when it contains multiple characteristic values. (20, 10, 5, 5)

## Lesson 9

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not understand the Cayley-Hamilton theorem.
4. I do not know how to use the Sylvester formula to represent a polynomial of a matrix. (20, 10, 5, 5)

## Lesson 10

1. I do not have enough time to finish my assignment.
2. I do not understand all the materials in the assignment.
3. I do not know how to determine the coefficients in the representation of a vector by the characteristic vectors of a nonsymmetric matrix. (20, 10, 6)

## APPENDIX B

Practice and Review Problems:

## Part I

Group A

1. The process of finding the combined value of two or more given numbers is called addition. For example, the combined value of 2 and 3 can be found by addition. The result of addition is called the sum. We find, by addition, that the combined value of 2 and 3 is 5, 5 is called the sum. Therefore, if John has 5 pears, Joan has 4 pears, and they have 9 pears altogether, --- is called the sum. (9)
2.  $2 + 4 = 6$  We do this by counting four numbers after 2, that is, 3,4,5,6. + is read plus. Therefore  $3 + 2 = ---$ . (5)
3.  $4 + 5 = ---$ . (9)
4.  $7 + 0 = ---$ . (7)
5. I have six books and you have three books, then we have altogether --- books. (9)
6. The sum of 4 and 2 is ---. (6)

Group E

1.  $6 + 2 = ---$ . (8)
2.  $7 + 1 = ---$ . (8)

3. The sum of 4 and 4 is ---. (8)
4. I have 4 blue pencils and 3 red pencils. I have --- pencils altogether. (7)
5. Mr. Smith worked 4 days last week, he works 5 days this week, he works --- days altogether. (9)

## Part II

### Group A

Problems 1,2,4,5,7,9,10,11,14,15 of group C.

### Group B

Problems 1,2,4,5,7,8,9,10,11,13,14,15 of group C

### Group C

1. When the numbers to be added have more than one digit, we add units to units, tens to tens, hundreds to hundreds, and so on. In performing the addition, add the units first, then the tens, then the hundreds, and so on. For example,  $13 + 25 = 38$  which comes from  $3 + 5 = 8$  (units to units),  $1 + 2 = 3$  (tens to tens),  $11 + 23 = \text{---}$ . (34)
2.  $42 + 37 = \text{---}$ . (79)
3.  $25 + 52 = \text{---}$ . (77)
4.  $35 + 10 = \text{---}$ . (45)
5.  $113 + 321 = \text{---}$ . (434)

6.  $212 + 531 = \text{---}$ . (743)
7.  $225 + 100 = \text{---}$ . (325)
8. I have 22 nickels, you have 37 nickels, we have --- nickels altogether. (59)
9. Mrs. Smith buys 135 eggs, Mrs. Jones buys 243 eggs, they buy --- eggs altogether. (378)
10.  $112344 + 362451 = \text{---}$ . (474795)
11.  $11 + 7 = \text{---}$ . (18)
12.  $132 + 5 = \text{---}$ . (137)
13.  $245 + 22 = \text{---}$ . (267)
14.  $125422 + 2173 = \text{---}$ . (127595)
15. The distance from Washington to Philadelphia is 127 miles, from Philadelphia to New York is 90 miles, then the distance from Washington to New York is --- miles. (217)

Group D

1.  $23 + 32 = \text{---}$ . (55)
2.  $133 + 331 = \text{---}$ . (464)
3.  $54321 + 12345 = \text{---}$ . (66666)
4.  $17358 + 1131 = \text{---}$ . (18489)
5.  $16215 + 1 = \text{---}$ . (16216)



Group E

1.  $313235 + 211321 = \text{---}.$  (524556)

2.  $13548 + 26200 = \text{---}.$  (39748)

3.  $1000 + 3 = \text{---}.$  (1003)

4.  $413234 + 73454 = \text{---}.$  (486688)

5.  $1314 + 8585 = \text{---}.$  (9899)

## Part III

Group A

Problems 1,2,4,5,7,9,10,13,14,15 of group C.

Group B

Problems 1,2,4,5,7,8,9,10,11,13,14,15, of group C.

Group C

1.  $2 + 3 + 4 = 9$  We do this by adding 2 to 3, and then their sum to 4, that is,  $2 + 3 = 5$ , then  $5 + 4 = 9$ ,  $1 + 3 + 5 = \text{---}.$  (9)

2.  $1 + 2 + 3 = \text{---}.$  (6)

3.  $1 + 2 + 2 + 3 = \text{---}.$  (8)

4. I have 2 apples, John has 3 apples, Mary has 3 apples, Paul has 1 apple, Anne has none. We have --- apples altogether. (9)

5. Similarly,  $25 + 31 + 12 = 68$  which comes from  $5 + 1 + 2 = 8$  (units to units),  $2 + 3 + 1 = 6$  (tens to tens)  $31 + 22 + 34 = \text{---}.$  (87)

6.  $41 + 14 + 34 = \text{---}.$  (89)

7.  $121 + 313 + 112 + 111 = \text{---}.$  (657)

8. I spend 125 dollars to buy a coat, 310 dollars to buy a TV set, 231 dollars to buy a sofa. I spend --- dollars altogether. (666)

9. Mr. Smith can write 1110 words every hour, Mr. Jones can write 1052 words every hour, Mr. Doe can write 1507 words every hour, they can write --- words every hour totally. (3669)

10.  $11 + 21 + 3 = 35$  which comes from  $1 + 1 + 3 = 5$  (units to units),  $1 + 2 = 3$  (tens to tens, notice here that only two numbers have the tens digits).  $31 + 14 + 2 = ---$ . (47)

11.  $13 + 2 + 41 = ---$ . (56)

12.  $1327 + 21 + 301 = ---$ . (1649)

13.  $1131 + 1 + 23 + 400 = ---$ . (1555)

14. I have 1000 dollars, my brother has 314 dollars, my sister has 22 dollars. We have --- dollars altogether. (1336)

15. I have three books, book A contains 1125 pages, book B contains 1030 pages, book C contains 833 pages. There are --- pages altogether. (2988)

#### Group D

1.  $22 + 33 + 10 + 14 = ---$ . (79)

2.  $313 + 100 + 201 + 304 = ---$ . (918)

3.  $1000 + 315 + 203 + 111 = ---$ . (1629)

4.  $113454 + 321 + 1201 + 10 + 2 = ---$ . (114988)

5.  $231624 + 132043 + 213022 = ---$ . (576689)

Group E

1.  $313 + 423 + 122 = \text{---}.$  (858)
2.  $111131 + 234142 + 203201 = \text{---}.$  (548484)
3.  $1428 + 100 + 10 + 1 = \text{---}.$  (1539)
4.  $813205 + 41323 + 2120 + 11 = \text{---}.$  (856659)
5.  $13212 + 21312 + 33133 = \text{---}.$  (67657)

## Part IV

Group A

Problems 1,3,4,6,7,8,9,10,13,15 of group C

Group B.

Problems 1,3,4,5,6,7,8,9,10,12,13,15, of group C

Group C

1.  $7 + 5 = 12$  We do this by counting five numbers after 7, that is, 8,9,10,11,12. Therefore  $6 + 8 = \text{---}.$  (14)
2.  $5 + 5 = \text{---}.$  (10)
3.  $9 + 8 = \text{---}.$  (17)
4. When the sum of two unit digits exceeds 10, we add 1 to the tens digit for every count of 10 in the unit digit.  $27 + 7 = 34$  which comes from  $7 + 7 = 14$  (units to units)  $2 + 1 = 3$  (tens to tens). Therefore,  $35 + 6 = \text{---}.$  (41)
5.  $58 + 7 = \text{---}.$  (65)
6.  $179 + 5 = \text{---}.$  (184)
7. I read 25 pages in the morning and read 8 pages in the after noon, I read totally --- pages today. (33)

8.  $23 + 4 + 5 = 32$  which comes from  $3 + 4 + 5 = 12$  (units to units),  $2 + 1 = 3$  (tens to tens). Then  $37 + 4 + 8 = \text{---}$ .(49)

9. John has 12 pencils, Mary has 8 pencils, Joan has 6 pencils, they have --- pencils altogether. (26)

10.  $25 + 17 = 42$  which comes from  $5 + 7 = 12$  (units to units),  $2 + 1 + 1 = 4$  (tens to tens).  $38 + 27 = \text{---}$ .(65)

11.  $48 + 29 = \text{---}$ .(77)

12.  $67 + 24 = \text{---}$ .(91)

13.  $41 + 35 + 17 = \text{---}$ .(93)

14.  $42 + 18 + 37 = \text{---}$ .(97)

15. The carpenters built 25 chairs day before yesterday, built 18 chairs yesterday, built 22 chairs today. They built --- chairs altogether. (65)

#### Group D

1.  $29 + 3 + 7 = \text{---}$ .(39)

2.  $38 + 49 = \text{---}$ .(87)

3.  $365 + 127 = \text{---}$ .(492)

4.  $17 + 18 + 19 = \text{---}$ .(54)

5.  $11 + 12 + 13 + 14 + 15 = \text{---}$ .(65)

#### Group E.

1.  $186 + 9 = \text{---}$ .(195)

2.  $39 + 57 = \text{---}$ .(96)

3.  $33 + 49 = \text{---}$ .(82)

4.  $13 + 19 + 22 = \text{---}$ .(54)

$$5. \quad 31 + 17 + 16 + 14 = \text{---}.(78)$$

### Part V

#### Group A

Problems 1,3,5,6,8,9,11,13,14,15 of group C

#### Group B

Problems 1,3,5,6,8,9,10,11,12,13,14,15 of group C

#### Group C

1.  $49 + 28 + 17 = 94$  which comes from  $9 + 8 + 7 = 24$  (units to units).  $4 + 2 + 1 + 2 = 9$  (tens to tens). Therefore,  $39 + 27 + 27 = \text{---}.(92)$

2.  $38 + 15 + 19 = \text{---}.(72)$

3.  $18 + 17 + 16 + 29 + 19 = \text{---}.(99)$

4. Mr. A. has 37 books, Mr. B has 19 books, Mr. C has 43 books, they have --- books altogether. (99)

5. I pay 7 dollars for a pen, 8 dollars for a book, 6 dollars for a shirt, 9 dollars for a pair of shoes, 8 dollars for a tennis racket. I spend --- dollars altogether, (38)

6.  $75 + 68 = 141$  which comes from  $3 + 8 = 1$  (units to units),  $7 + 6 + 1 = 14$  (tens to tens). Therefore,  $55 + 87 = \text{---}.(142)$

7.  $79 + 80 = \text{---}.(159)$

8.  $38 + 27 + 84 = \text{---}.(149)$

9.  $145 + 238 + 176 = 559$  which comes from  $5 + 8 + 6 = 19$  (units to units),  $4 + 3 + 7 + 1 = 14$  (tens to tens),  $1 + 2 + 1 + 1 = 5$  (hundreds to hundreds). Therefore,  $235 + 168 + 237 = \text{---}.(730)$

10.  $135 + 246 + 369 = \text{---}$ . (750)
11. I traveled 354 miles in January, 348 miles in February, 217 miles in March. I traveled --- miles totally. (919)
12.  $1728 + 459 + 362 + 70 = \text{---}$ . (2619)
13.  $173688 + 243197 + 3129 + 51301 + 4336 = \text{---}$ . (475651)
14. There are 24588 men, 29632 women and 14325 children in the city. There are --- persons altogether. (68536)
15. Mr. Smith made 156320 dollars in 1957, 230107 dollars in 1958, 186293 dollars in 1959, 231464 dollars in 1960. He makes --- dollars totally. (804184)

Group D

1.  $59 + 68 + 73 = \text{---}$ . (200)
2.  $351 + 248 + 269 = \text{---}$ . (868)
3.  $2632 + 731 + 749 + 27 = \text{---}$ . (4139)
4.  $15315 + 25326 + 77319 = \text{---}$ . (117960)
5.  $464286 + 211373 + 132310 + 101010 = \text{---}$ . (908961)

## APPENDIX C

A criterion to determine how a teaching program should be modified at a branching point is suggested in Chapter 5. This criterion is to maximize the "index number" which is a linear function of the student's accumulated grade. This linear function is computed by the "intrinsic linear operator" which is in the form  $\alpha + \beta X$ . A general discussion on this criterion is given below.

The intrinsic operator is chosen to be linear not only because this is the simplest form but also because (1) it is intuitively sufficient to have a linear dependent term to estimate the effect of the student's past performance on the modification of the teaching program, (2) in some cases, second or higher order non-linear operator would become an unrealistic criterion. This can be shown by plotting the index number against the student's grade  $X$  for different operators. Let  $\alpha_1 + \beta_1 X$ ,  $\alpha_2 + \beta_2 X$ ,  $\alpha_3 + \beta_3 X$  be three linear operators:

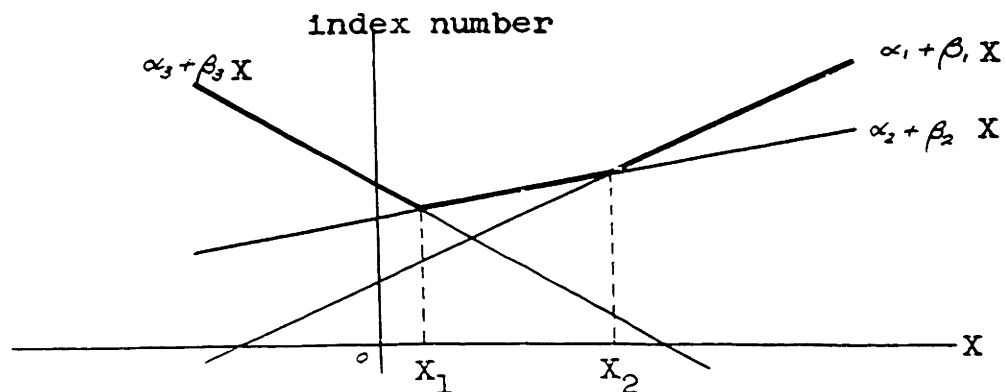
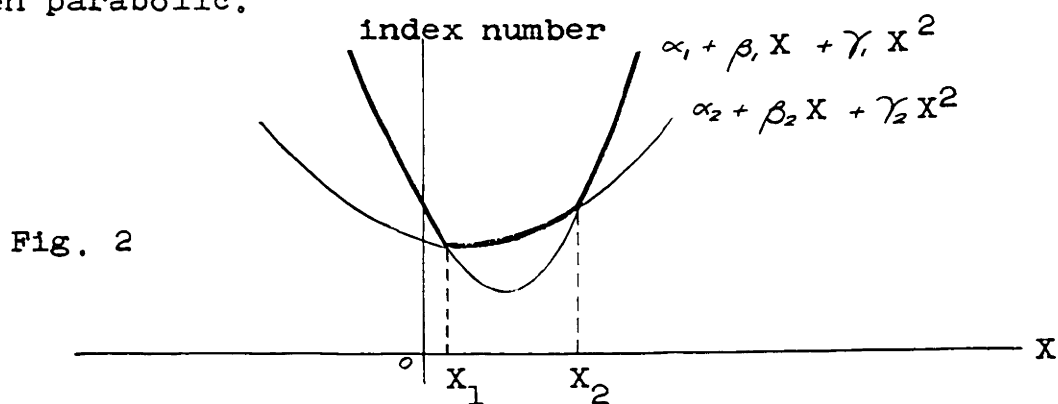


Fig. 1

The heavy lines show the highest attainable index number for different ranges of  $X$ . In this plot, it is clear that teaching program branches out depending on the students grade  $X$ . If  $X \geq X_2$ , the student is put into group I (corresponding to the linear operator  $\alpha_1 + \beta_1 X$ ); if  $X_1 \leq X < X_2$ , he is put into group II (corresponding to  $\alpha_2 + \beta_2 X$ ); and if  $X < X_1$ , he is put into group III (corresponding to  $\alpha_3 + \beta_3 X$ ). As mentioned before, when two operators give equal index numbers, we always break the tie by going to the higher group, e.g., in Fig. 1, when  $X = X_2$  we go to group I instead of group II.

As to the non-linear ones, let  $\alpha_1 + \beta_1 X + \gamma_1 X^2$ ,  $\alpha_2 + \beta_2 X + \gamma_2 X^2$  be the two operators. The relation between the index number and  $X$  is then parabolic.



Since, in general, two second order curves have two points of intersections, it turns out that these non-linear operators require that a student should be put into group I (corresponding to  $\alpha_1 + \beta_1 X + \gamma_1 X^2$ ) either if  $X > X_2$  (very good performance) or if  $X < X_1$  (very poor performance). This is obviously unreasonable to let the best and the worst students have the same teaching program while the medium students have a different one. Therefore, when non-linear operators are used, the grading system must be chosen in such a way that the student's



grade  $X$  will be in a range within which the parabolic curves have just one intersecting point.

As to the determination of the  $\alpha$ 's and  $\beta$ 's, we can either fix all  $\alpha$ 's at the beginning and adjust the  $\beta$ 's or vice versa. If we set all  $\alpha$ 's according to the practice offered by every group and choose one  $\beta$  as reference, we can find all other  $\beta$ 's by a graphical method:

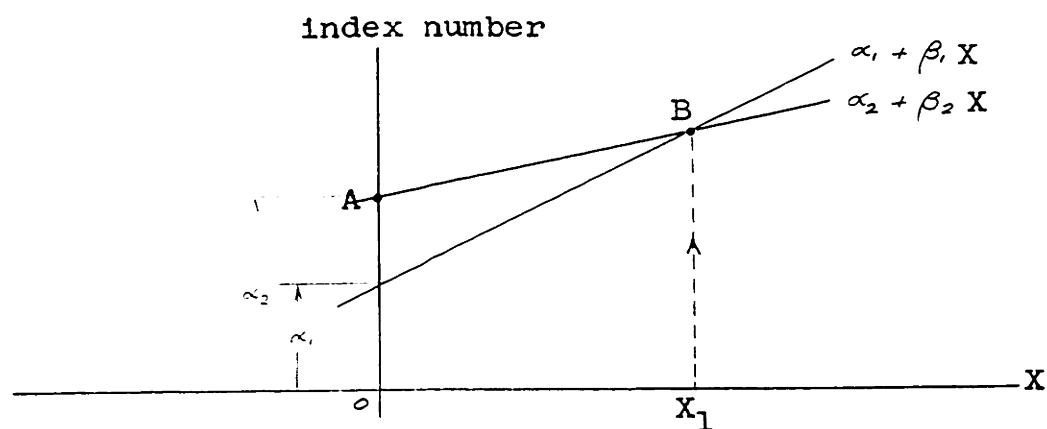


Fig. 3

Draw the straight line  $\alpha_1 + \beta_1 X$  where  $\alpha_1$  and  $\beta_1$  are pre-determined. Estimate the value of  $X$ , called it  $X_1$ , which will be the dividing grade between two categories of students.  $X = X_1$  corresponds to a point B on the straight line  $\alpha_1 + \beta_1 X$ . Join A, B, by a straight line whose slope would equal to the value of  $\beta_2$ . All the  $\beta$ 's can then be found in a similar manner. It is more expedient to do this algebraically. Equate

$$\alpha_1 + \beta_1 X = \alpha_2 + \beta_2 X \text{ at } X = X_1, \text{ gives } \beta_2 = \frac{1}{X_1} (\alpha_1 - \alpha_2) + \beta_1$$

The procedure is very similar when all  $\beta$ 's are fixed at the beginning. In Fig. 3, a straight line with a slope equal to  $\beta_2$  (given  $\beta_2$ ) is drawn through B. The intercept on the  $X = 0$  axis then equals to  $\alpha_2$ . Algebraically,

$$\alpha_2 = \alpha_1 + (\beta_2 - \beta_1) X_1$$

The values of the intrinsic linear operators labelled on the flow graph of Chapter 5 are chosen based on a set of fixed  $\alpha$ 's. Notice that the  $\alpha$ 's are equal to two times the number of practice problems in every group, also that the differences between  $\beta$ 's decreases as the teaching proceeds and the student accumulates higher X. If we fixed all  $\alpha$ 's instead, the intrinsic linear operators would be:

$$\begin{array}{rcccccc}
 & & & & & I_E = 46 - 3X \\
 II_A = 2X & II_B = 4 + X & II_C = 6 - X & II_D = 44 - 2X & II_E = 96 - 3X & \\
 III_A = 2X & III_B = 14 + X & III_C = 38 - X & III_D = 84 - 2X & III_E = 146 - 3X & \\
 IV_A = 2X & IV_B = 24 + X & IV_C = 68 - X & IV_D = 124 - 2X & IV_E = 206 - 3X & \\
 V_A = 2X & V_B = 34 + X & V_C = 98 - X & V_D = 163 - 2X & & 
 \end{array}$$

See here that the  $\beta$ 's are set equal to 2, 1, -1, -2, -3 respectively while the  $\alpha$ 's increase as teaching proceeds. Although this set of linear operators gives the same criterion on the modification of teaching program as the set labelled on the flow graph in Chapter 5 does, the increasing  $\alpha$ 's are not so satisfactory theoretically. Since it is not reasonable for two groups which offer the same amount of practice work to have greatly different value of  $\alpha$ 's e.g.,  $II_B$  has an  $\alpha$  equal to 4 and  $V_B$  has an  $\alpha$  equal to 34 while both of them require the student to do 12 practice problems out of 15. However, for a set of pre-determined  $\beta$ 's, the  $\alpha$ 's would usually turn out to be integer and thus make the computation easier.

## BIBLIOGRAPHY

1. Angell, G. W., and Troyer, M.E., "A new Self-Scoring Device for Improving Instruction", School and Society, 67:84-85, 1948.
2. Bernard, Guy G., Briggs, Leslie J., and Walker, Elbert S., "The improved Subject Matter Trainer", Technical Memorandum, Armament Systems Personnel Research Laboratory, Lowry Air Force Base, Colorado, April 1955.
3. Briggs, Leslie J., "Two More Self-Instructional Devices", Psychological Reports, 4: 671-676, 1958.
4. Bryan, Gleen L., and Rigney, Joseph W., "An Evaluation of a Method for Shipboard Training in Operations Knowledge", Technical Report No. 18, Department of Psychology, University of Southern California, September, 1956.
5. Bryan, Gleen L. and Schuster, Donald H., "The Effectiveness of Guidance and Explanation in Trouble Shooting Training", Technical Report No. 28, Electronics Personnel Research Group, University of Southern California, October 1958.
6. Crowder, Norman A., An Automatic Training Book on Number Systems, Volume I, Hoover Electronics Company, Timonium, Maryland, 1958.

7. Fry, Edward B., "Teaching Machine Dichtomy: Skinner Versus Pressey", Paper presented at the American Psychological Association Meeting, Cincinatti, September 1959.
8. Fry , Edward B., Bryan, Gleen L., and Rigney, Joseph W., "Teaching Machine: An Annotated Bibliography", Technical Report No. 28, November, 1959 p. 2-3.
9. Homme, Lloyd, "Research Notes", Automated Teaching Bulletin, Rheem Manufacturing Company, Volume I, No. 1, September 1959.
10. Holland, James G., "Teaching Machines: An Application of Principles from the Laboratory", paper presented before the 1959 Invitational Conference on Testing Problems sponsored by the Educational Testing Service.
11. Howard, Ronald A., "Machine-Aided Learning", Quarterly Progress Report No. 9, January 31, 1960, Computers Component and System Group, Electrical Engineering Department, M.I.T.
12. Irion, Arthur L., and Briggs, Leslie J., "Learning Task and Mode of Operation Variables in Use of the Subject Matter Trainer", Technical Report, Air Force Personnel and Training Research Center, Lowry Air Force Base, Colorado, October 1957.
13. Keislar, Evan, "Theoretical Aspects of Automatic Teaching", paper presented at the Western Psychological Association, San Diego, April 1959.
14. Perin, C. T., "The Effect of Delayed Reinforcement upon the Differentiation of Bar Response in White Rats", Journal of Experimental Psychology, 32, 95-109, 1943.

15. Peterson, J. C., "A new Device for Teaching, Testing and Research on Learning", Transactions of the Kansas Academy of Science, 33:41-47, 1930.
16. Porter, Douglas, "Teaching Machines", Harvard Graduate School of Education Bulletin, 3(1), March 1958.
17. Pressey, Sidney L., "A Simple Device for Teaching, Testing and Research in Learning", School and Society, 23:373-376, 1926.
18. Pressey, Sidney L., "A Machine for Automatic Teaching and Drill Material", School and Society, 25:549-552, 1927
19. Pressey, Sidney L., "A Third and Fourth Contribution Toward the Coming Industrial Revolution in Education", School and Society, 36:668-672, 1932.
20. Pressey, Sidney L., "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction", The Journal of Psychology, 29:417-447, 1950.
21. Ramo, Simon, "A New Technique of Education", Engineering and Science, 21:17-22, October, 1957.
22. Rath, Gustave J., Anderson, Nancy S., Brainers, R.C., "The IBM Research Center Teaching Machine Project", Automatic Teaching: The State of the Art, John Wiley and Sons, 1959, pp 117-130.

23. Skinner, B. F., "The Science of Learning and the Art of Teaching", The Harvard Educational Review, Volume 24, No. 2, Spring 1954.
24. Skinner, B. F., "Teaching Machines", Science, October 24, 1958, Volume 128, No. 3330, p. 969-977.
25. Weiner, P. K., "Proposed 'Automatic' Teaching Device", Institute of Radio Engineers Transaction on Education E-1:51-53, June 1, 1958.

Learning and the Art  
ional Review, Volume 24,

es", Science, October 24,  
69-977.

tic' Teaching Device",  
nsaction on Education





