SELF-ORGANIZING TEACHING SYSTEMS

EP-RR 20

S. KANEFF
A. N. VLADCOFF

December, 1968

Department of Engineering Physics
Research School of Physical Sciences
THE AUSTRALIAN NATIONAL UNIVERSITY
Canberra, A.C.T., Australia.

This book was published by ANU Press between 1965–1991.

This republication is part of the digitisation project being carried out by Scholarly Information Services/Library and ANU Press.

This project aims to make past scholarly works published by The Australian National University available to a global audience under its open-access policy.

SELF-ORGANIZING TEACHING SYSTEMS

by

S. KANEFF and A.N. VLADCOFF

December 1968

Publication EP-RR 20

Department of Engineering Physics
Research School of Physical Sciences
THE AUSTRALIAN NATIONAL UNIVERSITY
Canberra, A.C.T. Australia

CONTENTS

SUMMARY

Possible high level teaching systems using self-organizing principles are discussed. The 'evolutionary' character of a self-organizing system is used as a basis for a simple teaching system implemented on a digital computer. Deficiencies in the system suggest features of a more complex system which is being developed with the aim of providing a powerful research tool through which eventually high level systems may be developed. The general approach of finding constraints and action capabilities to produce a self-organizing system which can enhance its own capabilities through learning, is suggested.

1. INTRODUCTION

We are concerned here with the role of a machine in a high level teaching system. Hitherto, the machine itself has played an intellectually trivial part in attempts to produce 'teaching machines'. It is our view that this situation need not persist, and that the role of the machine can be considerably updated.

Given an adequate initial organization of information, a suitable physical (hardware) structure, and provided with sets of strategies, operating rules and decision criteria, together with the opportunity to communicate, gather and re-organize information, particularly the ability to produce and continually to update a model of the student(s), a mechanized system can be reasonably expected to:

(a) adapt its operation to take account of individual student capabilities and biases, selecting information for optimum presentation, determining the nature of difficulties, and instituting help sequences when student performance requires these;

(b) optimize its existing strategies, rules and decision criteria through experience with single students and with groups of students;

(c) act as an information store, allowing a student to ask for, and be supplied with, information which he desires or needs;

(d) provide adequate challenge and stimulation;

(e) generally provide a powerful measurement and control system for testing theories, approaches and strategies in a realistic environment, thereby providing information enabling the machine designers and educational research workers to enhance the system's capabilities and performance, ultimately enabling specification of a practical teaching system to be made.

The potential capability of a suitable machine structure to update its performance through the integrated results of interacting with hundreds of individual students should not be underestimated, as the process of improvement can be achieved extremely rapidly, unlike the corresponding process in a human teacher, whose experience must be developed over a long period, due to his limited information processing rate. What is very clearly necessary at this stage of machine development, however, is that the provision of new information on topics to be taught, the enrichment of structure and the provision of new strategies, rules and decision criteria, must come from the educational research worker and machine designer - the machine, by its own actions, can be expected to improve its performance, both short and long term, only by selecting strategies, rules and decision criteria (which are present), thereby improving and optimizing its performance and adapting to each student. In this sense, it might be said to 'learn' to be a better teacher, but (at the present stage of knowledge) it cannot be expected to 'invent' new strategies, rules and decision criteria, although, by acting within its structure and general organization, it might well produce sequences of interaction with a student which have not been anticipated by the machine designer. (Invention of new strategies need not be ruled out completely, however, because if given appropriate capabilities for re-structuring information, an opportunity to review information, and to seek further information by being able to 'ask questions', even this aspect might eventually be handled).

Essentially, it might be anticipated that a high level mechanized teaching system which can be devised in the light of present knowledge, would be an extremely powerful research tool in the hands of an educational research worker, for studying and defining the problems involved in developing a specification for a general purpose teaching system able to learn to improve its own performance. The place of the human teacher in such an environment would undoubtedly be as a manager of a powerful facility, guiding its application and updating its capabilities.

2. MECHANIZED TEACHING SYSTEMS

2.1 Inadequacies of Existing Systems

Predictions made (around 1960) that by the 1970's students would be spending their days in front of computerized teaching machines have not been realistic. However, the acquisition by an ever increasing number of educational establishments, including primary and secondary schools, of digital computers for general data processing requirements has provided a strong drive for employing such equipment in Computer Assisted Instruction Systems. At the present stage of development many further studies must be undertaken before confident recommendations can be made in regard to desirable systems. The need of a large amount of research seems urgent, particularly as large on-line time-shared educational data processing system are being installed in some Universities in the U.S.A. While a relatively small number of experimental multi-console systems may be warranted in order to develop this new concept in teaching, a great proliferation of expensive equipments likely to become rapidly obsolete does not seem desirable: (Rath (1967), mentions the proposed 700 console system at the University of Wisconsin and predictions from the University of Michigan that their system will have over 1,000 consoles by 1970).

A major weakness in present computer assisted systems lies in the inadequate means for communication between student and machine. This involves both the 'language' used for communication, which constrains the student to communicate in the machine's terms, as well as the effectiveness of the terminal equipment. In a multi-console system, the terminals themselves will certainly contribute a large proportion of the total cost, if facilities which are non trivial are to be provided - in this respect it is of interest to note an estimate by Bitzer et al (1967) that one central processer might be expected to process 3000-5000 terminals, and the report by Goldberg et al (1966) on plans for providing terminals (via television receivers and small keyboard consoles) in the homes of community groups of 5000-10,000 people, connected by co-axial cable to a central computer. In addition to technical and educational considerations, strong economic factors demand great care in solving satisfactorily every aspect of these terminal requirements.

A further limitation relates to the kind of material which can be effectively handled by existing mechanized systems. 'Drill and Practice' routines are processed comparatively well because the task and subject matter can be stated explicitly and can be programmed exhaustively and completely without detracting significantly from the value of the routines. Even here, however, a great amount of work is necessary to ensure that systems are satisfactory, but it appears the future of this kind of interaction is assured, for even at University level, there is much scope for the inculcation of factual knowledge in a limited context. On the other hand, the implementation of Tutorial and Dialogue Interactive Systems is severely hampered by the inability to process the unexpected response or the unforeseen difficulty.

More fundamentally, however, existing computer assisted approaches are structurally and functionally unable to take into account the student attributes and the character of the topic material in a sufficiently rich interactive manner - even if this aspect is overcome however, there remains a vast gap in knowledge concerning the inter-relationships between student and topic material parameters (including the problem of identifying those parameters which in fact are relevant) and the transformations necessary to achieve successful learning and understanding. Because of the inadequacy of existing theories of teaching and learning, it seems necessary to face the problem in an empirical, experimental manner in order to gather adequate information on which to base subsequent theories which would hopefully be useful in practice.

Computer Assisted Schemes usually select each following move on the basis of the immediate past response only, (although the general strategy of the interaction may have been, in the first instance, selected to match the student on the basis of general ability). This essentially minimally adaptive approach* is of extremely restricted use in higher level operation, as it takes little account of the overall performance over a period of time, and each of the prespecified branching routines, once selected, cannot take account of individual student differences which may apply over that routine.

There now seems little doubt that programmed instruction with or without mechanized implementation is a successful form of teaching within certain restrictions. Claims often made that it is successful because of its features of small step progressions and immediate knowledge of results, however, are difficult to substantiate with confidence, as other factors are invariably raised - the necessity for the programmer to first clarify his concepts and objectives, the great redundancy incorporated in the presentation of information, and so on - which contribute to the nature of the programme. The advisability of the small-step approach itself has been questioned often and has led to other approaches in which larger 'chunks' of information are presented - this seems particularly apt with bright students and the attempt to develop creativity (Crutchfield, 1965). These aspects need careful re-thinking continually, particularly as greater numbers of research workers and practising teachers

* See Section 2.2

take up writing programmes.

In the development of mechanized teaching systems, there has been a natural but regrettable tendency to accept existing digital computing systems and to employ 'state of the art' techniques in producing experimental teaching systems. Digital computers have not been evolved with the objective of solving teaching system problems and it is not surprising that a major mis-match exists in this approach. A consequence of this is, apart from the likelihood of inadequate systems resulting, that costs of systems, particularly when implemented in a half-baked manner, are likely to be needlessly great.

Undoubtedly there exists a pressing need to conduct intensive research into all of the various aspects entering into mechanized teaching systems - philosophical, sociological and technological.

2.2 The Machine as an Adaptive Control System

It has been observed (Stolurow, 1961) that the term 'adaptive' may be applied in a teaching system to indicate the capacity of the machine and its associated program to adjust, in some way, to the specific needs of the individual learner. We reserve the term 'minimally adaptive' for systems which have a finite (non-zero) minimum adaptive property present: that is, the machine should be able to adapt on the basis of the previous value, only, of a single measured parameter of a student, in producing the next item for presentation, in order to be called 'minimally adaptive'. (This parameter could, for example, be the time taken to respond, or the correctness or otherwise of a response). If more than one student parameter influences the succeeding item(s), or if the selection of items is made on the basis of 2 or more previous values of one or more student-parameters, the system is more than 'minimally adaptive', and it seems reasonable to designate it simply as adaptive. A memoryless system which is sensitive to 2 or more student parameters, or a system with memory operating on one or more student parameters, seems more than 'minimally adaptive' and will be termed adaptive. This use of the term 'adaptive' seems not inconsistent with that employed by Pask to describe his own developments of teaching systems. Further complexity is introduced when the factors, parameters or individual previous responses are themselves inter-related in a manner taken into consideration by the teaching system; then patterns of parameter variations and of responses may influence selection of items. Eventually the complexity of an adaptive system might be increased to the point where it would warrant the designation 'self-organizing' system, as introduced in Chapter 3.

Pask seems to have been the first to have produced (and understood the significance of) a device able to modify or adapt its presentation to match the student being taught, in more than just a trivial manner. His original work began with a system to teach keyboard skills (training punch-card operators) and has since extended to the teaching of such skills as making managerial decisions, the use of scientific inference in research, diagnosis and fault-detection in electronic equipment, and others. (Pask, 1957, 1958a, 1958b, 1960a, 1960b, 1961, 1962, 1965, 1967; Pask and Lewis, 1962). The principle used in all of these activities is that of including in the system an adaptive mechanism which interacts with the student, and allows the machine to learn appropriate student characteristics, which are used to modify the training routine according to these individual features. His devices have employed various forms of memory of past responses as well as memory of processed responses. Design of such a system involves a control-engineering problem, with the machine acting as the control mechanism to modify the student's behaviour (Pask, 1965).

Such a man/machine system can be compared to a homeostatic process, as it seeks an equilibrium defined in terms of accurate responses to the most difficult questions which the machine can offer (Beer, 1959). A dynamic equilibrium is set up at the completion of the task, with the machine presenting the most difficult questions, and the student correctly making responses to these questions. To achieve stability, the teaching machine makes trial moves which are initially unbiased; however, as interaction proceeds, the teaching machine learns about the student's behaviour pattern and, as a result of this, modifies its own strategy.

One requirement for stability of this interaction is that there should be a balance between the co-operative and the competitive strategies on the part of the machine. The machine takes the initiative. If the student is scoring well, it will compete, by giving less cue information and presenting more difficult items. This it will do until mistakes begin to appear, when the machine will start to co-operate by simplifying the problem (or perhaps giving a help sequence) with respect to the kind of mistake which was made. There is therefore a tight feedback between the two systems (student and machine). Because of the information transfer between the two systems, the adaptive teaching machine is similar to a participant in a dialogue.

Pask's work has been performed mainly by the construction of special purpose equipment, and appears to have been restricted principally to the training of various skills. Although highly successful, and certainly opening a vast new frontier of work, it has not attempted the problem of devising a universal (i.e. general purpose) system for teaching.

Systems which aim to be adaptive, in order to adjust to the specific needs of an individual student, must be able to discover the values of essential parameters - this can be achieved from his responses, and some form of memory in which to store these parameters (which need continual updating) appears necessary. The student and machine must be in constant interaction to achieve effective adaptation: the system of student plus machine may be regarded as a self-organizing system.

2.3 A Philosophy of Approach for Mechanized Teaching

The activity whereby a teacher (tutor) imparts knowledge and understanding of a topic to a student, is viewed as a communication, measurement and control process, in which the overall result is a transfer of particular information from an individually organized store of information within the teacher's brain, (including also information stored in books, and so on, to which the teacher has access and which he processes in his brain before imparting this information to the student), to a somewhat similar individually organized store of information within the student's brain. Because of hereditary and environmental factors, the organization and inter-relationships in the information within the two brains will differ to a degree which depends on these factors. Even the strategies for processing this information by the two individuals will probably differ, as will their attitudes, general levels of processing capability and so on.

The teacher, not knowing what actually transpires within the student's brain (and vice versa), can rely only on the communication media between them to provide information on the student's attitudes, capabilities, level of knowledge and understanding, and other attributes. In the case of a human teacher and student, the modes of communication are many - including speech, with a richness of intonation, inflexion and so on; visual, through the written, typed or printed word, or pictorial (diagram or picture); and the great variety of information which can be conveyed by general facial and body expressions. Whatever the mode, the teacher must rely on behavioural information in order to assess (measure) the student's internal state. He does this, probably largely subconsciously, and essentially builds up a model in his own mind of just what the student's attitudes and attributes are - he uses this model, which is revised (consciously or subconsciously) continually, and on it he bases his teaching strategies and item details, seeking to guide (control) the interaction in a direction producing the desired behaviour (goals) in relation to the topic being presented.

Insofar as he builds up a model of the student and can devise and select his strategies, decision rules and performance criteria, the teacher adapts his presentation to the needs and capabilities of the student. Further, over a comparatively long term, he learns, as a result of many encounters with various students, which approach is most effective for students with particular attributes. His success as a teacher depends in fact on how well he can match his approach to the student, and on how well he can select, measure and classify appropriate relevant student attitudes and attributes, and implement an appropriate set of strategies, decision rules and performance criteria suited to the teaching and subject matter goals.

Viewing the teacher/student interaction in the light of the above three paragraphs, suggests the model on which the herein described systems are based. As illustrated in Figure 2.1, the Controller (experimenter or system designer), has direct access to the student and to the teaching system, his aim being to match operation of the overall system to produce effective teaching-learning interactions, employing appropriate communication media between teaching system and student. The store of information within the teaching system is organized by the Controller on the basis of his own conceptual understanding of the topic material - the

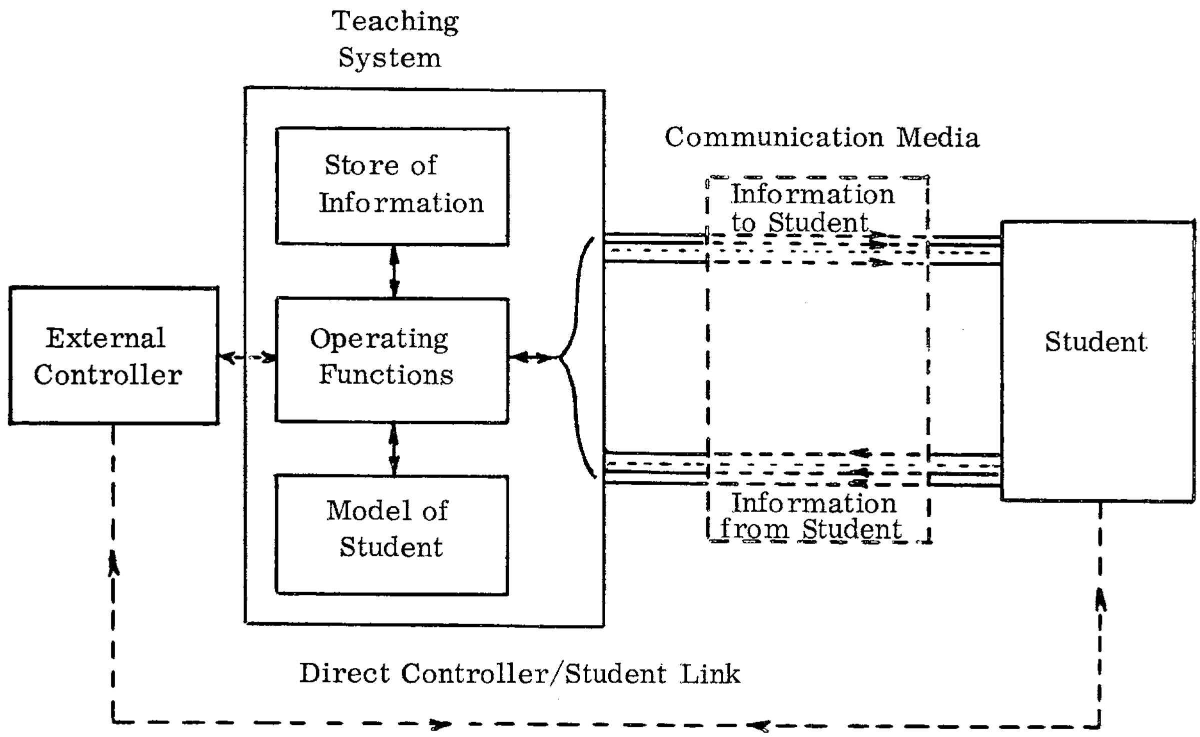


Figure 2.1 Elementary Schematic of Controller/Student/Teaching System.

model of the student reflects this information store in structure so that, hopefully, when the teaching-learning interactions have concluded, it will also reflect the information store content in detail. Necessarily, the store of information in the teaching system can be only a very crude version of the organization within the Controller's brain, but nevertheless, this may be an advance on no organization at all; similar remarks apply to the relation between the student and his model incorporated in the teaching system.

The model of the student contains other details dealing with his general capabilities and attitudes - these attributes or parameters are taken into account, together with the properties of the information and general teaching goals, strategies, decision rules, and performance criteria (all inserted by the Controller) in order to produce teaching-learning interactions well matched to the student's attributes. A feature of this approach is the necessary complementary relationship between the teaching system structure and the functions which must be inserted into the structure to make it operate successfully. The task of the controller, apart from initially devising the necessary structure, organizing the information, and inserting the functions required, is to monitor behaviour with the object of improving the structure and including new functions. The teaching system itself is expected to be able to optimize on the existing functions and to learn from experience which functions apply in particular cases - it must be provided with operating strategies to do this.

Further clarification of the relationship between the information store and the student model is necessary. It is noted that (almost) intuitively a human teacher can sum up a student's capabilities - with a little thought, he can give an opinion on the student's cognitive style, his attitudes, and other less readily definable qualities. The teacher is able to do this because of his own attributes, some of which have been developed through a great deal of training and experience. If the teacher desires to discover directly the attributes of a total stranger, he can do this essentially in two ways - he may wait to discover them gradually through a (relatively) long personal acquaintance, or he can set out deliberately and quickly to discover them by test. Thus, inferential or deductive abilities may be discovered by testing with respectively inferential or deductive items of information. By organizing a topic to be taught not only in terms of concepts and their inter-relations, but also with each individual item for presentation to the student devised for a purpose (e.g. to discover inferential, deductive, and so on, abilities) and 'tagged' accordingly, including details of the concept itself, the responses of the student can be assessed in relation to the 'tagged' characteristics which they carry and thus, over a reasonable length sequence of interaction, a rich store of student attributes can be assembled. The only effective limit imposed by this approach is the ingenuity of the controller in identifying appropriate student parameters and devising (and suitably labelling) items to develop and discover the student's attributes in relation to these parameters. Thus the information store is organized in terms of concepts and their inter-relationships with all items for presentation, duly 'tagged' with the kind of parameters they contain, or are used to discover as attributes in the student. The student model, then, contains essentially a replica of the information store in terms of its conceptual organization, so that the details of the student performance to the particular topic can be 'filled in' as interaction proceeds - that is, his degree of completion for each aspect of the topic material is recorded and continually updated. Another aspect of the model contains a structure enabling a continual updating with regard to other student attributes as deduced from the scoring to the 'tagged items', as well as quantities such as error rates, response time, and more general attributes such as motivation, persistence and so on.

Further student attributes may be derived from his attitude and behaviour (for example, from his speed of response) and from his general performance to the strategies, rules and criteria incorporated within the teaching strategy. All of these are factors for inclusion, with suitable tags, within the student model.

It is of fundamental importance to our approach that <u>the continually updated values of student attributes within the student model</u>, (together with information which can be derived from these attributes) <u>provide the only possible quantities or characteristics of the student which can be measured and/or controlled by the teaching system</u> (whether a human teacher or machine): consequently the richness and relevance of the parameters incorporated in the student model essentially determine the level of adaptation possible - this is true whatever the nature of the system and can serve as a profitable basis on which to compare different teaching systems. The need for discovery or identification and quantisation of useful student attributes is therefore of paramount importance to the future progress of high level Adaptive Teaching Systems.

In order to achieve the necessary teaching goals and to adapt to the particular student, a human teacher is guided by certain principles and strategies (which he may or may not have clearly enunciated), which he has acquired through training and experience, to present and explain his material, and to adjust in accordance with the responses of the student. The teacher is characterized by extreme flexibility in relating his approach to the subject matter and student attributes, so that he can usually achieve his purpose. It is necessary to provide a corresponding capability in the mechanized high level teaching system, so that appropriate adaptation can be achieved on the basis of the parameters stored in the student model - in this case, however, the principles, strategies and other features cannot remain vague and not enunciated, (unless a completely random and impracticable approach is employed to gather, then organize, information). The stage of development of mechanization of cognitive processes does not at present allow invention or creativity in a practical sense (e.g. Amarel, 1966; Minsky, 1961; Solomonoff, 1966) and the controller seems obliged to provide to the machine an adequate set of strategies, decision rules, performance criteria, as well as teaching goals and subject matter goals.

In view of the partly intuitive nature of the human teaching approach, specification of particular strategies, rules, and so on, is a difficult assignment. It seems in the very nature of people that they learn, even when there are no apparent specific goals of learning indicated, or no pre-specified approach to reach these goals available. Effective teaching merely accelerates and directs this process by providing a suitable environment in which the learning goals are more readily reached (the environment is essentially attained or modified by the features of measurement and control of student attributes in relation to the subject material, and the goals, strategies, decision rules and performance criteria applied). How this may be done is a further field for fundamental research, no less important than the determining of the appropriate quantities on which measurement and control can be based, as discussed earlier.

Even the best human teachers are not perfect in always selecting the most appropriate goals, strategies, rules or criteria: moreover, because of limited processing ability where large masses of data are concerned, teachers do not remember detailed facts about individual students, only general principles and attributes. Where a teacher interacts with large numbers of students over a period, it takes him a long time to process the large amount of data resulting from these interactions, into operating rules based on this experience. Even in the human case, invention is a rare occurrence, and a good teacher's approach is based on training and experience. Again, because of human difficulty in handling the complex decision making and large amounts of data necessary in situations involving multivariable processes, it is unlikely that a teacher will produce an optimum learning environment for every given student (even while tutoring to individual students), because both factors taken into account in making decisions and the data itself are first simplified, and there is little time to even try to optimize. Consequently, a human teacher is likely to learn to teach reasonably well relatively rapidly, as he can adapt quickly to individual needs as a result of his general human capabilities, but will find that it takes a relatively long experience to make further improvements when these depend on experimentation and optimization over a large number of trials. (Indeed, many human teachers seem not to bother to take this last step).

A mechanized adaptive system on the other hand, can have powerful information processing features in that it can handle and store large masses of data, and can engage tirelessly in numerous experimentation trials and can perform optimization studies taking account of all data. The results can be organized and employed to improve in an optimum way, the teaching and subject matter goals, strategies, decision rules, and performance criteria, all within a comparatively short time. Moreover, complete records of internal operation and details of machine-student interaction may be stored and/or printed out.

It is readily apparent that human capabilities, if combined with the information handling powers of the adaptive machine, can make a very useful and effective co-operative system for Educational Research, with a central feature of the continual improvement in performance which can be achieved as a result of experiences and modification. Because of the extreme present lack of knowledge on every aspect of the problem, an approach is favoured on the basis of an overall general structure which allows details relating to its various features to be incorporated to the extent of present knowledge, but with the capability of adding further details and even refinement to the structure without the necessity of major modifications. The approach adopted, in the absence of adequate theories of learning, must be to a degree, empirical, with experimentation directed deliberately to the gathering of information necessary to refine the structure and enrich the details of the functions within the structure.

2.4 Functional Requirements

Figure 2.2 illustrates diagrammatically the general requirements in more detail than Figure 2.1. From information provided by an appropriate set of stored student parameters (obtained initially from previous interactions, by standard tests, or inserted by the experimenter) the teaching system should be able to select or generate items from an appropriately organized information store relating to the topic, based on appropriate teaching goals, subject matter goals and on a student parameter dependent set of strategies, decision rules and performance criteria, and present these items, suitably displayed, to the student. The student's responses are to be assessed in relation to the performance criteria, and the results used to update the student parameters, provide information to adjust strategies, rules and criteria. Study of the pattern of responses in relation to the items presented, the goals, strategies, rules and criteria should allow the isolation of difficulties by the generation of

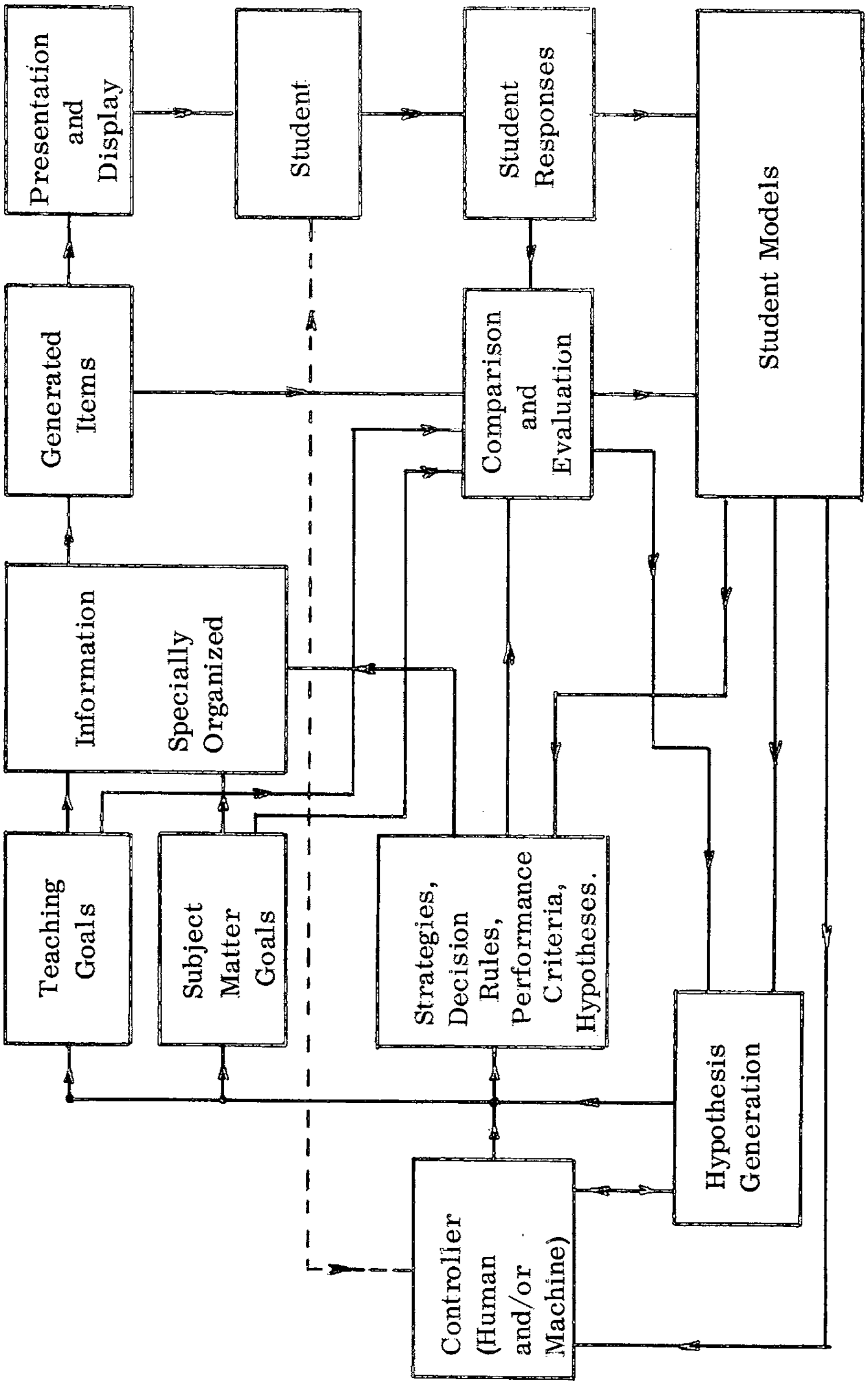


Figure 2.2 General Functional Requirements.

hypotheses which are then used to assemble sequences of items designed to test if the particular hypothesis indeed applies. The confirmed hypotheses must then cause appropriate help sequences to be organized to resolve the difficulty. The system should be able to improve its own performance by modifying its goals, strategies, rules and criteria as a result of interactions with individual students and with groups of students, (Kaneff, 1967).

Consideration on how the structures discussed above might be implemented raises fundamental problems in every aspect. Among these, educational objectives need to be defined, information must be appropriately organized and represented, adequate meaning must be given to 'performance' and how this is to be measured, the nature of the student model must be determined, adequate strategies and decision criteria must be evolved, and effective communication established. Interaction should involve not only the machine, in a sense, directing the learning process, but the student should be able to seek information which he considers necessary, and the machine should be able to supply it. The system should be capable of continual updating of its performance capabilities.

2.5 Complexity and Self-Organization

On consideration of the kind of general purpose teaching system structure which would meet the philosophy of approach and requirements outlined, it is clear that a quite complex system emerges. The system must be multivariable with complex feedbacks, and, because it must adapt to individual student needs and peculiarities - that is, it must consider some goals, strategies, rules, and criteria as more probable than others for a given student, not allowing any of these to be dropped completely in case they are useful for a student of different attributes, or the same student during periods of different motivation, etc. - the system must be basically probabilistic. An approach in which there are clear cut decisions on the above factors, is inferior to the probabilistic approach because of the consequent lack of variety in items for presentation and in coping with the same student over spells of disinterest and variable attributes.

Self-organizing systems appear relevant in this environment: they are systems which can learn and increase their behavioural repertoire through the capability of building up a model of their environment by learning.

3. SELF-ORGANIZING SYSTEMS

The term 'self-organizing system' has been used in the literature in a variety of ways, usually with reference to adaptive and learning systems, particularly to describe very complex processes such as those to be found in various living organisms. We employ the term here in the sense of Beer (1967) and Ashby (1962), to describe very complex, probabilistic systems which (a) can change their organization from a state of independent parts to a state of dependent parts, or (b) can improve their organization to fulfill a certain purpose or goal, or both (a) and (b).

As implied by the above, no system of itself can be self-organizing - it can do so only as part of a larger system. Thus, a teacher cannot improve his performance if he is in complete isolation; but a larger system of teacher and students in a real world environment, can be a self-organizing system. The theory of self-organizing systems has received a great deal of attention in the literature over the past decade, with reference to order, entropy and other concepts: we will refrain from invoking these formalizations here - see, for example, Beer (1967), Von Foerster (1960). Instead, it will be convenient to discuss informally the essential ingredients of a self-organizing system and to relate these to possible teaching systems.

At the outset, it is necessary to draw attention to certain possible misleading aspects of natural self-organizing systems. When we observe the end result of a set of natural processes, we are apt to interpret this subjectively as a goal. This viewpoint, although sometimes a useful one, can mislead by diverting attention from the underlying processes involved. Thus, if we have a cluster of gas molecules in an otherwise empty closed box, we notice that in due course, the molecules are spread sensibly uniformly throughout the volume of the box. From this and similar observations we deduce 'Laws' of Nature or goals relating to entropy and changes in order. We can view the processes differently: given an environment which imposes a set of constraints, and given particles of gas with certain action capabilities (e.g. the ability to move when acted on by forces), the whole system moves probabilistically and inevitably to a situation in which the gas particles are uniformly dispersed. The end result appears as a goal to an observer.

3.1 Relevance to Teaching Systems

Beer (1967), suggests that in the process of a self-organizing system moving to its 'goal', 'entropic' movement carries the system inevitably toward a pattern which guarantees an equilibrium between the system and its environment. When the observer has defined a set of goals as appropriate to his purpose, and has spotlighted an entropic drift which conforms to those needs, he labels the system as self-organizing, and refers to the changes taking place as evidence of Control. Further, the degree of control exerted is dependent on the amount of effective information available to the system.

This picture is similar to the view taken of the general purpose teaching system, which can achieve measurement and control of the student attributes only on the basis of the information available about them. Moreover, the teaching system is probabilistic, predictive and incorporates learning, all of which are identifiable features of Self-Organizing Systems. Consequently it seems appropriate to approach the design of the general purpose teaching system from the viewpoint of certain aspects of the principles of Self-Organization, which can be helpful in the consideration of systems with complex multivariable control functions.

The problem of designing a Self-Organizing System to take advantage of 'entropic drifts' reduces to one of choosing and maintaining a suitable environment within which the entropic drift can naturally and inevitably take place. For this process to occur, the utmost information must be allowed to flow in order to achieve effective control. Foremost requirement, however, is that a goal or set of goals, understood as much as possible, must be viewed as the inevitable outcome of a set of environmental constraints and action capabilities on the part of the system. It is the constraints and action capabilities which the designer must be ingenious enough to determine in order to produce a successful Self-Organizing System.

The entropic process which drives self-organization is homeostatic as Beer (1967) points out, but the variety generator which gives proliferated states from which successful patterns are selected is conditioned from outside by restricting total possible states; variety is diminished by feedback precluding other states, and by a learning mechanism which biases the alleged randomness of the mutations. Similarities in the teaching system to the above factors are readily apparent.

3.2 An Elementary 'Evolutionary' Teaching System

To attempt implementation of the principles so far put forward, an elementary teaching system has been simulated on a digital computer (Vladcoff,1968).

The development of the teaching machine structure outlined below has been based on the principle of achieving features of self-organization by generating 'structures' (by the application of certain rules), and selecting those which survive in a specified environment. (By 'structure' is meant any ordered entity). Self-organization in this sense, is that outlined in (a) in the first paragraph of this Chapter 3. This, in a very elementary sense, is a basis of the process of evolution. The motive in approaching the problem in this way, has been that self-organizing features can be obtained by modelling this evolutionary process which has been successful in producing complex and intelligent creatures.

In applying this philosophy of approach to a teaching machine system, the 'structures' generated are items for presentation to the student. The student fills the role of an environment, in that he is the criterion against which items are tested. Those items which are useful (in that they enhance learning) are retained, while those serving no useful purpose, in this way 'die off'. This has been achieved by adjusting the probabilities of selection of corresponding items.

It is pointed out that in the simulation actually carried out, the generation of items has been restricted to the class of meaningful items. The organization which occurs in the system during its 'evolution' in an environment (a student) is then one of continued improvement in presenting items <u>matched</u> to the needs of the student (a form of adaptation). Furthermore, because the environment is a changing one (due to an increase in student proficiency generally), the system must survive under these conditions, by issuing items which will continue to teach. This calls for certain memory requirements.

The organization which results in the teaching system can be plotted in n-dimensional space. This gives a certain path of development of the machine structure as interaction proceeds. However, (as in the case of evolution, where different trees of organization are formed depending on the particular environment), a different path of organization will result for each individual student.

The system developed illustrates certain principles of generation of structures from more basic components. This has application of great importance in teaching systems; for example, in generating items and strategies not foreseen by the programmer.

Although illustrating certain principles, the system has obvious limitations as discussed later.

3.2.1 Essential Features of a Simple Evolutionary Model

The process to be modelled is that of a homogeneous set of parts becoming organized in different ways, depending on the particular environment. The basic essentials for such a process can be listed as follows (see Figure 3.1).

(a) Basic raw material. This constitutes building blocks which are to be structured in certain ways. An example is a raw material used in nature, namely the amino acids. These basic elements are organized into various proteins during the growth of any cell: the actual protein formed depends on the type and arrangement of the various amino acids (of which there are some twenty known different ones acting as building blocks, used in the protein structure.

(b) A rule of operation. This is a rule for the generation of organizations. The rule is actually a function for ordering the raw material into complex structures. It may be said that the genes of the chromosomes within the nucleus of the living cell function as such rules of operation.

(c) A criterion. The purpose of this is to act as a measure against which the structures or organizations formed can be tested. Basically, it is a criterion for distinguishing between 'good' and 'bad' organizations. In nature, the environment itself provides such a criterion of acceptable organization. It behaves as a selective mechanism in that the 'good' organizations live on and reproduce, while those which do not meet the requirements specified by the environment will die. It is important to note that what is a good or a bad organization is relative, depending on

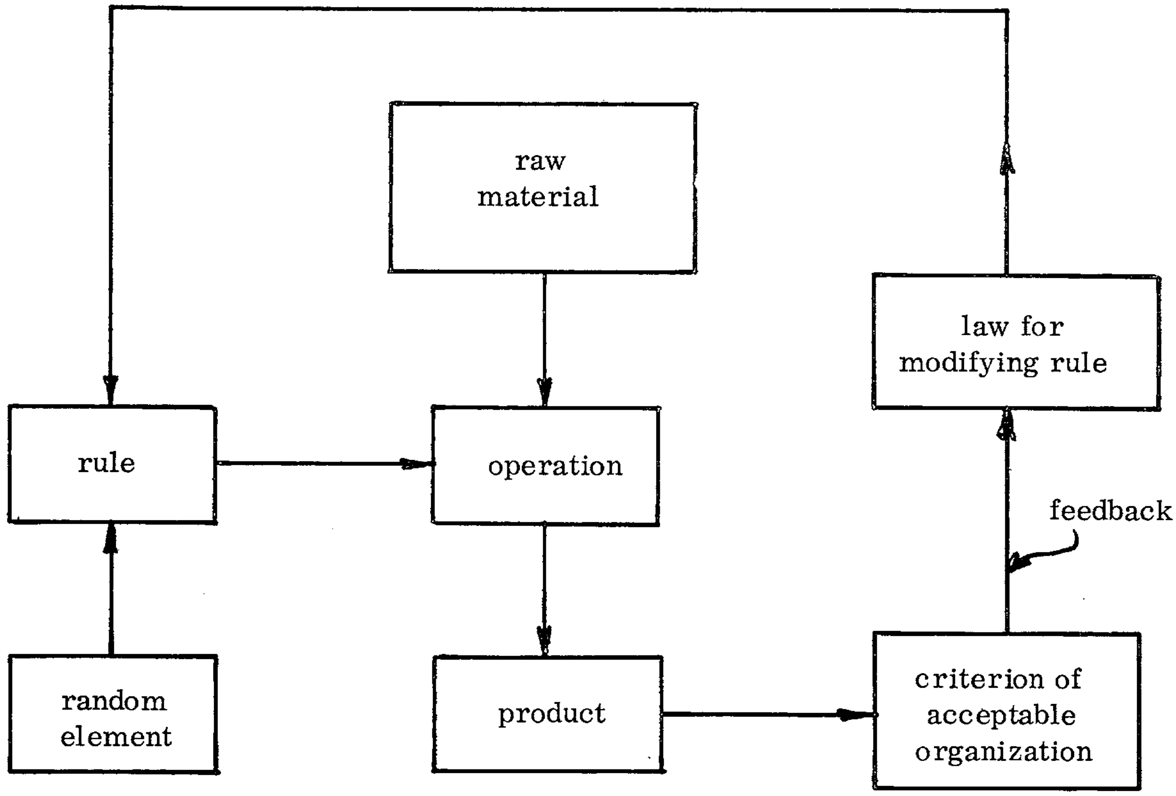


Figure 3.1 Simple 'evolutionary' model.

the actual environment (Ashby, 1962). What has survived in our particular environment must therefore be necessarily 'good'. But this is not to say that this organization will be acceptable in any other environment.

A criterion is necessary in this model because the sense in which self-organization is used here, is that of 'parts unconnected' changing to 'parts connected'. In such a case, there is no guarantee that the organizations formed will be acceptable, and hence the necessity to have some measuring stick against which they can be tested.

(d) A random generator. This produces the constituent which Von Foerster (1960) refers to as noise, and which Ashby (1952b) calls random information. It is best illustrated by the way it operates in nature. Here the noise arises from random gene mutations, which are a key to new creations in the evolutionary process. If such mutations were not present, the evolutionary system would merely settle down to a dynamic equilibrium, and new creations would cease. Ashby's homeostat (Ashby, 1952a) uses random information in the sense that the various feedback factors are adjusted at random, if the system ever becomes unstable.

Thus for any process of self-organization, both order (present in the form of rules of operation) and noise (from a random generator) are essential constituents.

The four factors listed above have been summarized by the simplified model of evolution shown in Figure 3.1. This model shows products being generated on the basis of certain rules, together with the intervention of a random element. After being tested against the environment, some will be retained, and others rejected. Because the products which are retained carry within them the rules for generating new products (in the form of the genes), there is present in effect, a factor which modifies the existing rules. This factor is represented in the model by the block labelled "law for modifying rule".

3.2.2 Description of the Teaching System

The system based on the model outlined, is illustrated in Figure 3.2. Briefly, the following are the essential features.

(a) Store of concepts

The set (C_1, C_2, ..., C_i) of concepts to be taught is the first portion of the lesson, and corresponds to the first sub-goal. Sub-goal 2 is then achieved after sub-goal 1, and so on until sub-goal n has been reached, which signifies the completion of the whole lesson. The criterion for the completion of each sub-goal is that there should be answered, in sequence, a prespecified number of the most complex type of question in that sub-goal. It is emphasized that the sense in which the word 'concept' is used throughout this chapter is as a single aspect of a skill or lesson to be mastered (the example of section 3.2.4 makes this clearer). Corresponding to each 'concept' is an item designed to instruct or test on this particular aspect of the material in the lesson.

The ordering of the sub-goals in this way is representative of initial structure which has been stored in the machine. This can be done on the basis of experience as to which order of arrangement is most suitable. However, there is not as yet any definite ordering of concepts within each sub-goal. This ordering or structuring is developed during an initial stage of interaction with a group of students, which prepares the machine for teaching.

This arrangement, it is seen, leads to a compromise between initial structure inserted, and structure developed as interaction proceeds. Such a correct balance is necessary in order that the 'teacher' is not too rigid in behaviour, yet at the same time has some constraint on the order in which material should be taught.

The grouping of concepts in each sub-goal is so

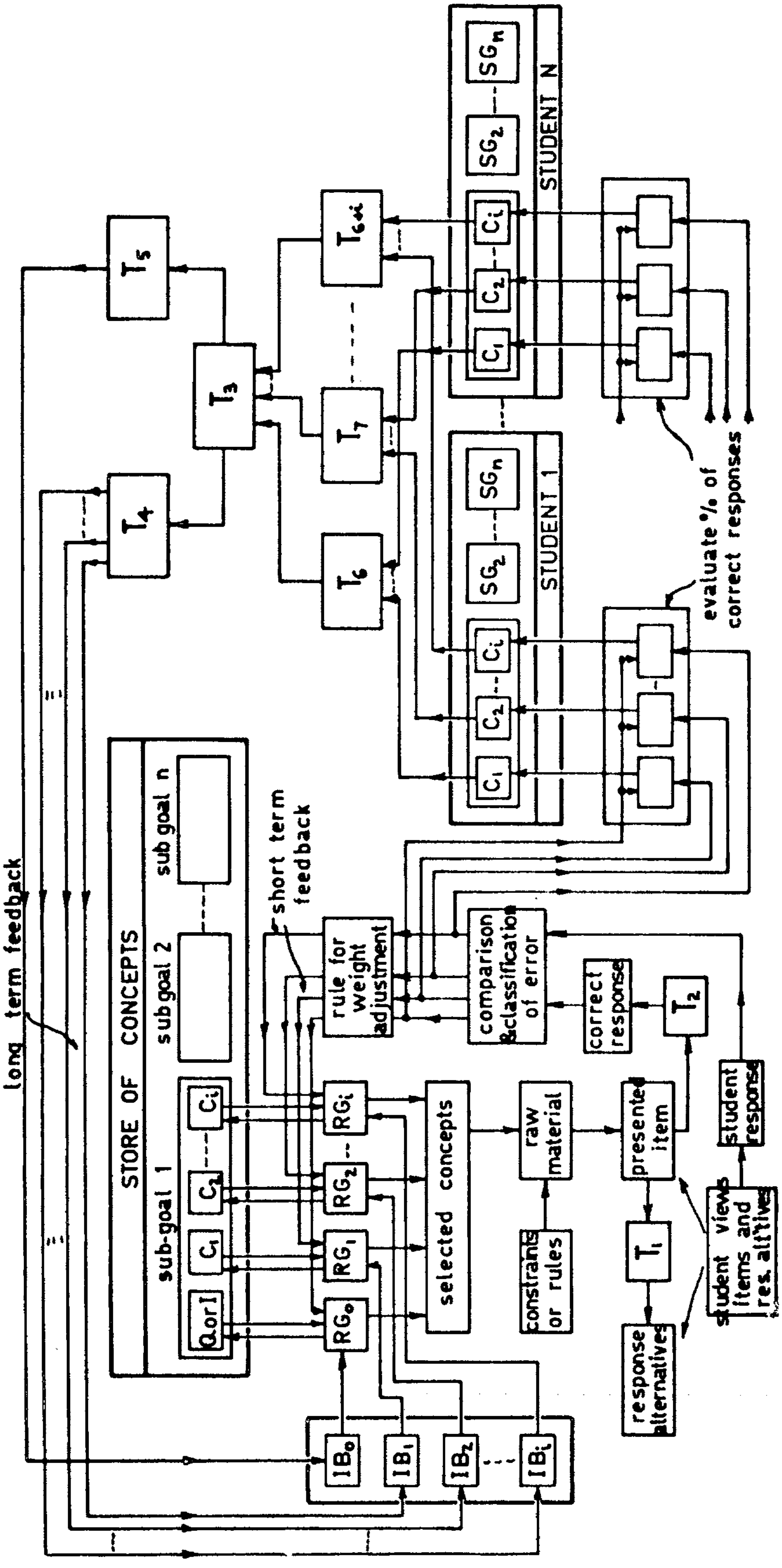


Figure 3.2 Overall structure of preliminary development.

arranged that irrespective of which combination is selected from the i concepts, a sensible item will result from their combination. There will therefore be,

$$\binom{i}{1} + \binom{i}{2} + \cdots + \binom{i}{i}$$

possible combinations, where $\binom{n}{x}$ is the number of ways of choosing x items from n .

(b) Random Component

Each of the blocks labelled RG in Figure 3.2, is a random generator which (without any further influence), would produce and present items taken completely at random from the information relevant to the sub-goal. This randomness is carried to the extent that the item itself may be a question or an instruction (refer to block, Q or I). Although a feature of the system is to allow this freedom of presentation of items in a sub-goal, it is however, advisable to place some constraint on the order of their presentation at the beginning of the interaction. This is achieved by an initial bias (IB) on each random generator (RG). The relative values of the biases placed on these random generators are assessed from human experience in teaching that particular subject matter, and will therefore be somewhere near their optimum values. For example, it may be known from experience that it is generally better to present C_2 before C_3 . In such a case, IB_2 and IB_3 will be set so that the probability of C_2 appearing before C_3 is enhanced. However, the actual optimization of all initial biases is carried out during an initial stage of operation, where the system is allowed to interact with a group of students.

(c) The Student Model

A rather simple student model was incorporated in this preliminary development, as shown in Figure 3.2. As can be seen, the performance of each student is assessed and stored in a separate model, which records his performance to all concepts in the curriculum, the storage being based on the percentage of correct responses.

The performance of groups to each concept, is analyzed (after the completion of each group) as indicated by T_6, T_7, $\cdots$, T_{6+i} , i being the number of concepts in the sub-goal in question. On a basis of percentage of correct responses, typical statistical analyses may be, (see Figure 3.3),

T_6 : Peak occurring at 85%
(i.e. most students score 85% correct responses to C_1).

T_7 : Peak at 62% .

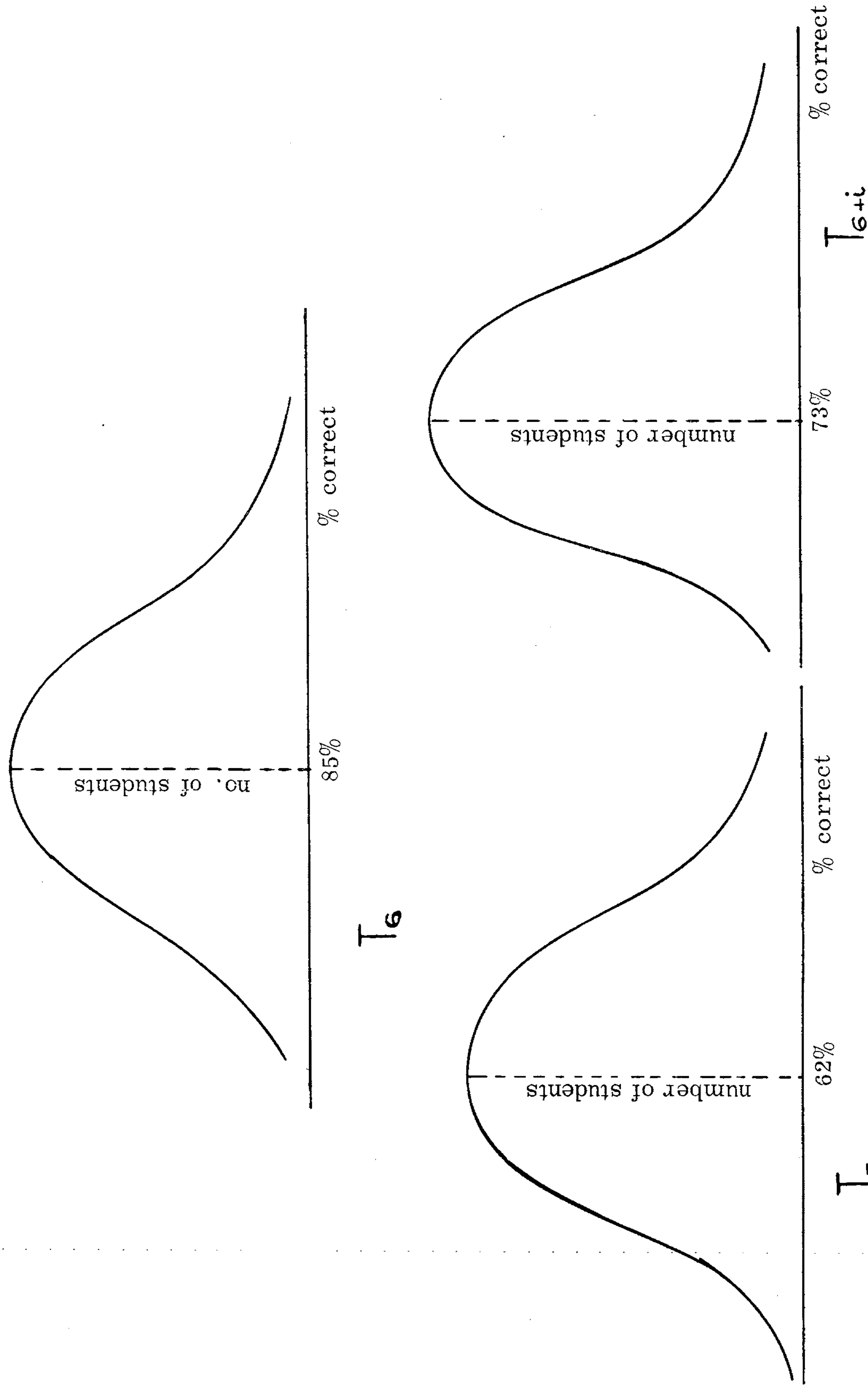


Figure 3.3 Typical student error curves.

T_{6+i} : Peak occurring at 73% .

Such statistical data is used in adjusting the initial biases IB_o, IB_1, $\cdots$, IB_i through the long term feedback indicated in Figure 3.2 .

3.2.3 Operation of the System

The two aspects of most interest in the operation are: the method of selection of items during normal interaction with a student, and the long term adjustment of initial biases, (which is performed after a group of students has passed through the lesson).

(a) Method of Selection of Items

The selection of an item from the concept store is governed by two factors.

(i) The long term performance of the group, as incorporated in the setting of the initial biases on the random generators.

(ii) The individual student performance as interaction proceeds. This is incorporated by the short term feedback (see Figure 3.2), which directly biases the random generators for making a choice of concept or group of concepts. (Without the bias on the random generators from this short term feedback, together with the initial bias, items would be issued purely at random).

Thus, item selection is made primarily on individual student performance, but has also ingredients of group characteristics, as well as a random component. For example, suppose that on previous group performance, it had been found that most students experience difficulty if items incorporating both concepts C_1 and C_2 are presented before items incorporating C_1 alone. Then the initial strategy of the machine would be towards presentation of C_1 prior to C_1 and C_2 together (IB_1, IB_2, $\cdots$, IB_i achieve this biasing). However, as there is a random component present, variations from this procedure are possible, which if successful for any particular student, are reinforced through the short term feedback. In this way, although item presentation is partially constrained, sufficient freedom is allowed to cope with student behaviours removed from the norm.

Once the concept or group of concepts has been selected, this is transformed to an actual item. (In the actual simulation made, a simple syntax which constructed sentences, was used). A comparison of student response with correct response is made, which leads to a weight adjustment (through the short term feedback) of the random generators, thereby directly influencing the next item presentation. The student's performance is also updated after each response by evaluating a new percentage of correct responses.

A typical rule for weight adjustment is to increase the probability of selecting a concept which has been poorly understood. Other rules consider the weight adjustment to be made to the random generator governing the proportion between questions and instructions. For example, a good performance will generally give a bias towards a higher frequency of questions, and so proceed through the task more quickly, while a poorer performance will lead to the presentation of more items of instructional value. These rules, used to transform student response into an appropriate weight adjustment, may not be optimum when the machine is first used (e.g. student errors may be too severely treated by presenting too many items of instructions when an error occurs. Or the rule used may not allow for sufficient remedial items to be presented for a certain error). However, such rules can be readily modified from the experience gained in using the system.

The questions presented by the machine may require a constructed response, or a response from various response alternatives on the part of the student. If the question chosen requires a response of the latter type, the transformation T_1 (see Figure 3.2) enables the relevant response alternatives to be evaluated. If an item is chosen which involves n concepts, then in the most general case, 2^n response alternatives are required in order that a choice from one of these by the student, will give sufficient information as to where the lack of understanding lies. However, in the course of interaction, the system, having assessed which of these n concepts has been understood, will limit the number of response alternatives so that only those which lead to new information about the student are presented. As an example, consider a lesson designed to distinguish green-round, green-square, red-round and red-square objects. If the system is satisfied that the student has learned colour, then it need not always present four response alternatives, and in such a case, two are sufficient. The exact form of T_1 depends on the particular lesson being taught.

(b) Adjustment of Initial Biases

When each student of the group has completed the lesson, a statistical analysis of the group performance is made, on the basis of performance to each of the concepts (see Figure 3.2, blocks T_6, T_7, $\cdots$, T_{6+i}). Operator T_3 then analyzes this statistical data, and by means of operators T_4 and T_5 (which are actually rules for weight adjustment), the initial biases on the random generators are adjusted. As an example, for the figures T_6, T_7, T_{6+i} quoted previously, a typical rule is to adjust the initial biases so that,

(i) The order of presentation of C_1, C_2, C_i is more likely to be:

$$C_1, C_i, C_2$$

and (ii) A higher percentage of instructions to questions occurs when C_2 is selected, than when C_1 or C_i is selected.

Again, such rules need to be modified from experience in using the system. For example, the setting of T_5 may be such that if less than 90% correct responses are received, a higher frequency of instructions is presented. What this setting should actually be, is a problem in Educational Research.

As this feedback is made only after a group of students has interacted, it is termed a 'long term feedback', (as distinct from any adjustments which may be made to the system during the course of interaction with any one particular student). The long term adjustment can also be made after each group of 20 or 30 students has taken the lesson.

3.2.4 Application to a Specific Problem

The problem chosen to illustrate operation of the system by a continuous generation of items, is the apparently simple task of teaching to tell the time. Only a brief outline of the relevant concepts involved in the task is given.

(a) On the Implication of the Term 'Concept'

The term 'concept' has been used in a variety of ways in the literature, and 'working definitions' have appeared. For example, Piaget (1957) considers a concept as an explanatory rule, or law, by which a relation between two or more events may be described. (The explanatory rules need not be classificatory). A concept can be represented by mapping the set of arguments onto the set of values of the function, and if only the two sets are known, the function might be found by induction.

Bruner et al (1956) state, "There are those who argue that a concept, psychologically, is defined by the common elements shared by an array of objects and that arriving at a concept inductively is much like 'arriving at' a composite photograph by superimposing instances on a common photographic plate until all that is idiosyncratic is washed out and all that is common emerges. A second school of thought holds that a concept is not the common elements in an array, but rather is a relational thing, a relationship between constituent part processes. We submit that such a controversy is relatively fruitless. We have found it more meaningful to regard a concept as a network of sign-significate inferences by which one goes beyond a set of observed criterial properties exhibited by an object or event to the class identity of the object or event in question, and thence to additional inferences about other unobserved properties of the object or event. We see an object that is red, shiny, and roundish and infer that it is an apple; we are then enabled to infer further that 'if it is an apple, it is also edible, juicy, will rot if left unrefrigerated, etc.' The working definition of a concept is the network of inferences that are or may be set into play by an act of categorization."

While these and other 'working definitions' are interesting and to a degree, informative, the task of actually handling information requires a somewhat less complex approach, at least initially, which can simplify a programmer's problem in organizing the information. Consequently, it is not intended here to argue on the merits of various points of view, but to employ the term 'concept' as describing aspects of a skill or lesson to be mastered in a practically useful manner, particularly to act as 'building blocks', appropriately inter-related, in forming higher level 'concepts'. This usage is, in many respects, similar to that employed by other writers. The test of whether certain concepts have been grasped by a student is made in terms of activities he can perform or problems he can solve.

(b) Basic Concepts in Time Teaching

The following are some aspects of behaviour to be mastered in the task of learning to tell the time.

(i) The distinction between the two hands.

(ii) The recognition of numbers 1-12.

(iii) The recognition of numbers 5-55.

(iv) The recognition of a hand pointing to a number.

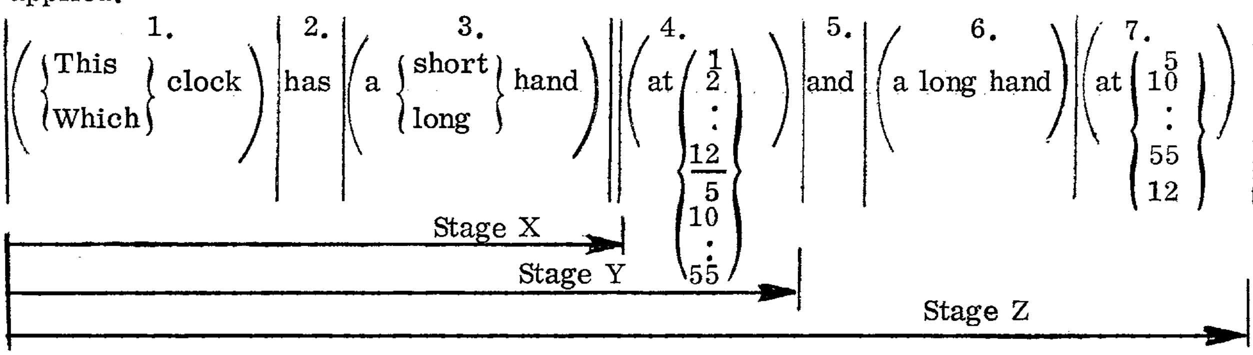
(v) The recognition of which clock has a certain hour reading.

(vi) The recognition of which clock has a certain hour-minute reading.

(There are of course other aspects of understanding involved, which are considered in more detail at a later stage).

(c) Syntax for the Formation of Sentences

To carry out initial experiments, the following simple structure illustrates the formation of a variety of sentences, depending on the concepts or combination of concepts selected by the random generators, with appropriate biases applied.



The rules for the formation of sentences (constraints) in this example are:

(i) If 'short' is selected in column 3, then a selection from column 4 is made <u>only</u> from the numbers 1-12.

However, if 'long' is selected, then a selection from any of the numbers 12 or 5-55 can be made in column 4.

(ii) If 'short' is selected, the system can form an item up to stage X, Y or Z. But if 'long' is selected in column 3, then the program is such that stage Y is the highest order item formed.

(The actual choice and construction of items is considered in (e)(i)).

With a simple structure as the above, items of the following type can be generated:-

This clock has a short hand.
Which clock has a short hand? (2 items)

This clock has a short hand at $\left\{\begin{matrix}1\\ \vdots\\ 12\end{matrix}\right\}$
Which clock has a short hand at $\left\{\begin{matrix}1\\ \vdots\\ 12\end{matrix}\right\}$? (24 items)

This clock has a long hand at 12.
Which clock has a long hand at 12? (2 items)

This clock has a long hand at $\left\{\begin{matrix}5\\ 10\\ \vdots\\ 55\end{matrix}\right\}$
Which clock has a long hand at $\left\{\begin{matrix}5\\ 10\\ \vdots\\ 55\end{matrix}\right\}$? (22 items)

This clock has a short hand at $\left\{\begin{matrix}1\\ 2\\ \vdots\\ 12\end{matrix}\right\}$ and a long hand at $\left\{\begin{matrix}5\\ \vdots\\ 55\\ 12\end{matrix}\right\}$
Which clock has a short hand at $\left\{\begin{matrix}1\\ 2\\ \vdots\\ 12\end{matrix}\right\}$ and a long hand at $\left\{\begin{matrix}5\\ 10\\ \vdots\\ 55\\ 12\end{matrix}\right\}$? (288 items)

Note: In this latter case, if the selection of '12' is made by the long hand, so obtaining,

"Which clock has a short hand at 6, and a long hand at 12?",

then the program translates this to,

"Which clock shows 6 o'clock?"

The simple structure in this case produces a variety of 338 different items.

(d) Correlation of the Problem with the General System

The relationship of the formation of items, to the block diagram of Figure 3.2 is as follows:-

(i) Block 'Q or I' $= \left\{ \begin{matrix} \text{Which} \\ \text{This} \end{matrix} \right\}$

This is controlled by RG_o with an initial arbitrary setting $IB_0 = 0.4$ to 0.6 (questions to instructions).

This arbitrary value was evaluated on the a priori basis of beginning a lesson by presenting, in general, one or two instructions followed by a question, that is, a ratio of 1.5 to 1 (which initial approach to the lesson is varied from group experience).

(ii) Block $C_1 \equiv$ Concept of length $= \left\{ \begin{matrix} \text{long} \\ \text{short} \end{matrix} \right\}$

The selection here is controlled by RG_1, with an initial setting favouring presentation of C_1 before any presentations involving more than a single concept.

(iii) Block $C_2 \equiv$ Recognition of nos. 1-12 .

Selection here is controlled by RG_2 . In this case, the system is biased such that C_2 alone does not appear (although it can be modified to to do so). C_2 appears only in conjunction with C_1, giving items of $C_1 C_2$ conceptual content (which relate to identifying to which of the numbers 1-12 , a hand points).

(iv) Block $C_3 \equiv$ Recognition of nos. 5-55 .

Selection is controlled by RG_3, with the bias again set so that C_3 alone does not appear. As indicated earlier, items which are meaningless are not generated, and so no items with content C_2 C_3 are formed. However, items which do appear are, for example, those with content C_1 C_3 (which relate to identifying to which of the numbers 5-55, the long hand points).

(v) Block C_4 $\equiv$ Concept of telling time in hours.

Selection is controlled by RG_4, with a setting of IB_4 such that items containing C_4 are not likely to be issued early in the lesson (although the possibility still exists and if the move is 'successful', reinforcement occurs, so allowing the possibility of the good student passing very quickly through the task).

(vi) Block C_5 $\equiv$ Concept of telling the time in hours and minutes.

Here the probability of such items occurring is controlled by RG_5, with a bias on IB_5 such that these are least likely to occur early in the lesson.

However, it is pointed out that all the initial biases indicated (although initially set by the programmer), are subject to modification from experience with groups, and so can be set at a more suitable value.

(e) Simulation of the System

(i) Programming Features

The basic flow diagram of the system is shown in Figure 3.5, while the method of generation of items is indicated in Figure 3.4 .

Figure 3.4 shows the key words which need to be selected in the formation of items. This selection is made by the parameters indicated using the following coding.

K = Parameter for formation of items involving C_1 (initial bias is KSC).

(The SC term following parameters refers to an initial bias).

L = Parameter for setting a question or an instruction (initial bias is LSC).

MS = Parameter for formation of items containing C_2 (initial bias is MSSC).

ML = Parameter for formation of items containing C_3 (initial bias is MLSC).

MH = Parameter for formation of items containing C_4 (initial bias is MHSC).

MF = Parameter for formation of items containing C_5 (initial bias is MFSC).

Figure 3.4 shows certain initial biases (set before interaction - for example, the probability of an item of instruction appearing is 0.6 , while that of a question is 0.4 . These initial biases, together with the weight adjustment through feedback, assign values of probability to the parameters. (In Figure 3.5, T = True, F = False).

As an example, consider that MF has been selected, with the value of L causing a question to be asked, MS indicating to concentrate on number 5, and ML indicating to concentrate on number 35 . Then the item presentation selected is,

"Which clock has a short hand at 5, and a long hand at 35" ?

(refer Figure 3.4).

The actual weight adjustment is achieved by an incrementation of variables for corresponding correct responses. For example, to adjust the random generator used as part of the selection mechanism for instructions and questions, an incrementation is made in L .

The details for this are as follows:-

A random number,

```
R = L * RANDF (1.)
```

is generated, (FORTRAN CODE).

which results in a random number between 0 and L (the initial value of L is 20).

Then if R-LSC is > 0, a question is chosen. Otherwise the item presented is an instruction. (LSC = 12 was the initial bias used, thereby giving a 0.6 probability of instructions).

An incrementation in L each time a correct response is received, then increases the probability of questions (there now being a greater chance of R-LSC > 0). The degree of incrementation in L can be varied, which

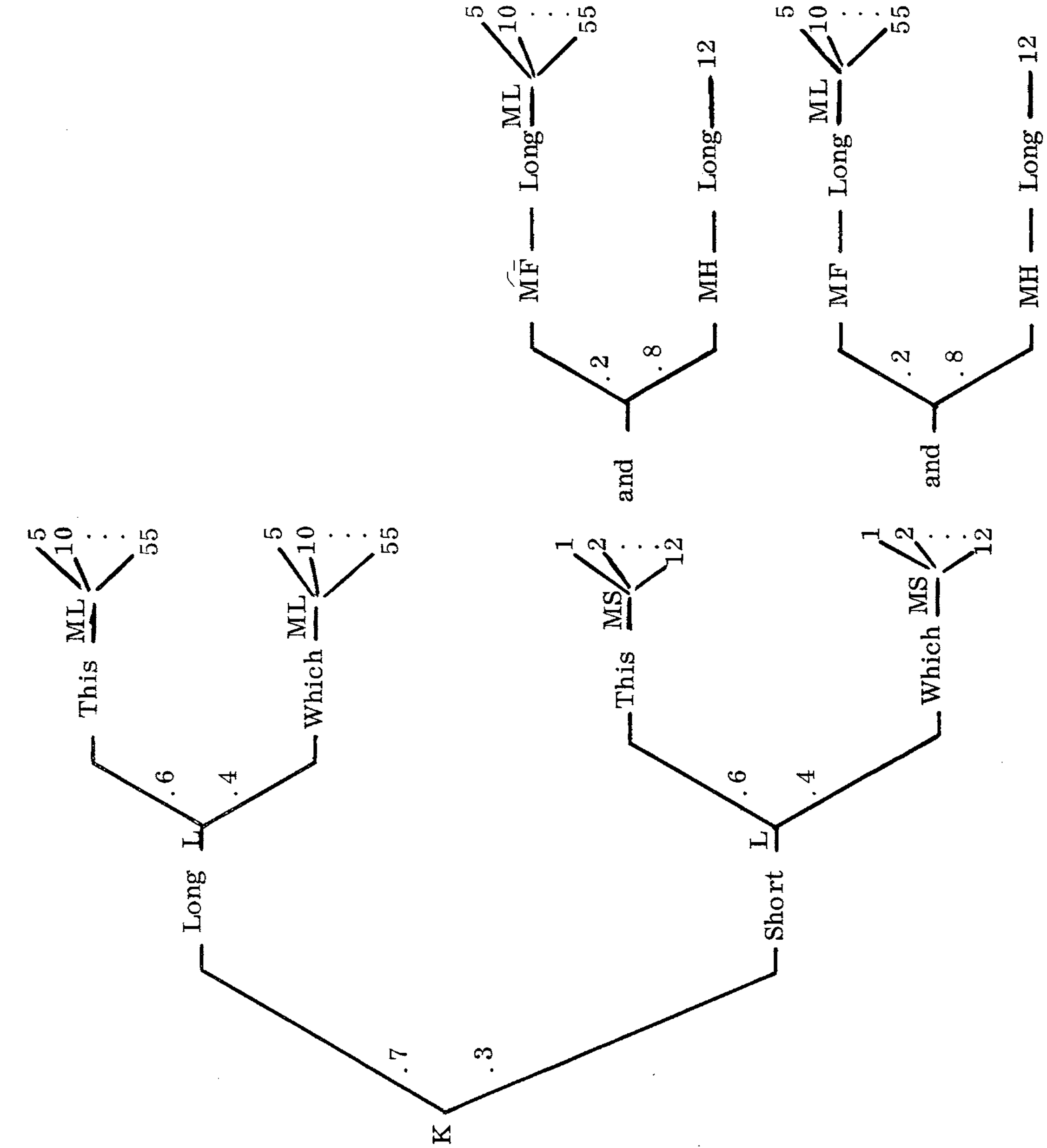


Figure 3.4 Method of item generation.

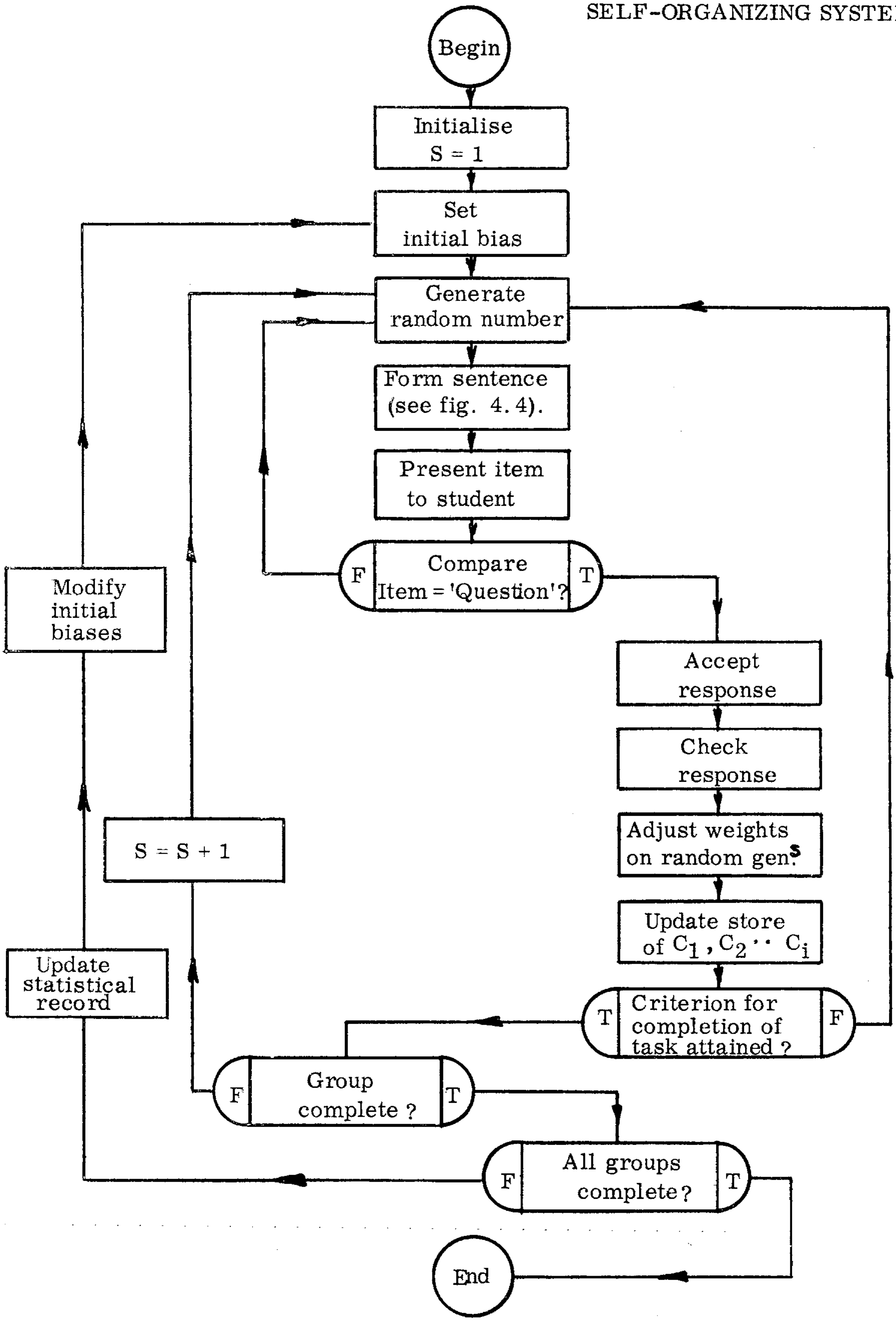


Figure 3.5 Outline of overall operation.

in effect is a modification of the rule for weight adjustment in the 'short term feedback' loop.

A similar form of weight adjustment is used for the other variables.

The interaction ceases when a specified number (in this exercise chosen between 5 and 20) of the most complex questions has been answered correctly.

(ii) Control of the Visual Display by Transformation T_1

The following indicates the type of analysis carried out by T_1 in forming various response alternatives for visual display. A complete analysis has been made for each question which can be presented, but a sample only is presented here to indicate the nature of the transformations involved.

Suppose the item presented is,

"Which clock has a short hand at (6) ?"

(Such an item presentation involves C_1 and C_2, that is, concepts of length, and recognition of numbers 1-12 - the actual number of the hour appearing is controlled by C_2, which considers the student's performance to each of the individual numbers).

When an instruction such as the above appears, T_1 performs the following:-

(1) One of the clocks (chosen at random) has a short hand placed at 6 (or a fraction past the 6, because fractional hours are also considered). The long hand is set at the corresponding correct position in relation to this short hand.

(2) The second clock is then incorrectly set, at a position close to the 6 (between 5 and 6), and the long hand set accordingly. The setting of the second clock in this way makes the selection of the correct response more difficult.

(3) Other alternative settings of the second clock face are:

Long hand between 6 and 7, (which then fixes the position of the short hand); or

The short hand placed at random, with the long hand then placed correctly in relation to it.

It can be seen that the choice of the 'incorrect' response alternative is a form of cue information. A choice of one of the above response alternatives is made on the simple performance figure indicated earlier.

(iii) Practical Aspects of Simulation

Owing to the inaccessibility of a graphic display unit, response alternatives, such as the above, were generated on a standard line printer (IBM # 1403). The long hand setting was made to the nearest 5 minutes, with the short hand taking up positions either half-way, quarter way, or three-quarter way between any two hour-settings. For initial tests, and in the absence of suitable communication facilities, the experimenter acted as the response mechanism, by playing alternately, roles of a 'good', 'medium', and 'poor' student to test the workability of the system.

A desirable feature apparent in the interaction was a yery rapid progression through the material (with few instructions) in the case of the 'good' student; with a correspondingly slower progression (with more items of instruction), in the case of a poorer student.

However, certain limitations were apparent, as discussed in the next section.

3.2.5 Discussion

(1) The system outlined has only a simple memory, and limited capabilities of processing information from student records. However, the feature of limited learning from experience, does take the system out of the class of purely adaptive systems. It can, perhaps, be compared with the lower forms of organism, which show a small degree of learning from experience (due principally to a limited memory), as well as adapting to their environmental conditions. The system outlined closely resembles this type of behaviour.

(2) Certain features of self-organization can be noted. For example,

(a) 'Differentiation' occurs along different pathways, depending on the particular interaction with the environment (that is, adaptation occurs).

(b) There is a certain improvement in performance with experience due to the long term feedback from group analysis.

(c) The system is probabilistic, in that it is not possible to define explicitly the next state of the system at any point in time. It is pointed out, however, that with time (as the operation proceeds) the system behaviour becomes more and more predictable.

Although properties of self-organization can be observed, they are of course at an elementary level, due mainly to insufficient complexity of structure and a limited memory to carry forward useful strategies learnt.

(3) Similarities between the system developed, and those of Pask's adaptive systems (for training of skills), can be noted (Pask, 1958, 1960, 1961). For example,

(i) Presentation is governed by student performance figures, with an emphasis on presentation of items least well handled;

and (ii) A corresponding matching or adaptation of the system to the participant.

There are, however, basic differences, including:

(a) A more flexible presentation (about a norm) is achieved by inclusion of a random element acting on previously-acquired student behaviour figures. This has the advantage of allowing a very rapid progress through the lesson by a good student (due to useful 'mutations' which arise). As far as we can ascertain, Pask has not considered adjustments on the basis of group behaviour.

(b) No allowance is made in this model for variation of rate of presentation. Pask does include such a feature, which is essential for attaining competence in certain skills. In the system outlined, the emphasis is more on rate of progress through the lesson concepts, rather than the physical rate of presentation of items. (The change of emphasis here is due to the different nature of the task considered).

(4) A limitation of this particular model is that selection of items is based principally on the individual student's most recent performance, and to some extent, on the integrated effect of the group. The deficiency here is that insufficient consideration is taken of the integrated effect of a student's overall performance.

(5) It was found in stability studies with the system, that the feedback factors used (i.e. the 'rules for weight adjustment') were rather critical, (e.g. it was possible for the system to over-compensate by presenting too many items of a given concept, if there had been a weakness in this concept earlier). A further cause of such instability is the absence of a more general student performance figure based on his overall performance.

(6) A further limitation of the memory is that in storing only a 'percentage correct' figure, certain vital aspects of student behaviour are never recorded.

(7) While the system is modelled in a manner resembling an evolutionary process, there is a difference in that unfavourable presentations (for a particular student) are not completely removed, but only have their probability of occurrence reduced. The reason for this is that an unfavourable presentation in one environment (i.e. for one student), may not be so for another, (cf. Ashby, 1962). It is therefore necessary to retain the possibility of such an 'unfavourable' presentation being issued.

(8) The behaviour of the system can be likened to a passive network, in that its behaviour is controlled and shaped completely by the particular environment. In other words, the presentation it gives depends on the environment in which it operates. (The system is in fact a function transforming student behaviour into item presentations). The limitation here is that the system cannot assume an active role, and do something positive if a student difficulty arises.

(9) Teaching to tell the time, as handled by the teaching system discussed in this chapter, seems somewhat like the training of a skill. The system outlined appears to be more suited to tasks which involve the training of skills, (e.g. keyboard skills), because of the limited extent to which items can be logically presented to teach tasks requiring deep understanding.

(10) The principle of generating items in a somewhat random manner and testing against an environment, offers the possibility of generating items which may not have been pre-selected by the programmer. To achieve this, a syntax is required which can operate on a set of item parameters, so producing sentences.

(11) Experimentation with the setting of operators T_4 and T_5 offers the possibility for Educational Research in the two problems (among others):

(a) The error rate problem, that is,

"Is there an optimum error rate, and if so, what is this rate?"

This problem can be investigated by studying performances of groups with different threshold values applied to T_5.

(b) The programming problem, that is,

"What is the correct order of presentation of material in a lesson?"

Again, this problem may be investigated from group studies, with each group assigned to different operators T_4 for influencing the order of presentation on a basis of past student performance.

3.2.6 Conclusions

The simple model presented has been based on constituents of the mechanism of evolution. The mechanism consists of a rule which orders raw material into different structures, together with a criterion of acceptable organization against which the structure can be tested.

This model has been applied in the design of a teaching machine structure which organizes its behaviour to suit the particular environment in which it operates (that is, a student). The structure itself is a compromise between pre-programmed behaviour, and behaviour learnt in a specialized environment. Before the machine is ready to teach, it undergoes a training period by interacting with groups. In the operating cycle, when the machine behaves as a teacher, its behaviour at first has a degree of uncertainty, but this is reduced as the machine - student interaction proceeds. The selection of material by the device is based on a particular student's most recent performance, and on the long term integrated performance of a group of students. This limited memory reduces the effectiveness of the system in several ways, notably in limiting the amount of learning from experience. In this respect, it resembles a lower organism.

The principle under which the system operates gives rise to the possibility of generating items (or strategies) not pre-programmed before interaction. This kind of system may be useful for research in solving problems which arise in mechanized instruction.

The above development has highlighted certain essential requirements in designing a more general purpose system which can be used for a wide variety of tasks requiring understanding in their attainment, as well as practice.

4. FURTHER WORK

The simple model discussed in the previous chapter has shown various limitations and suggested lines of approach for further work. The emphasis is on a general structure which can be 'shaped' by special purpose programs to handle particular topics. A great deal of flexibility is introduced in allowing various functions (strategies, decision rules, performance criteria) as well as new information, to be added at any stage, so permitting diverse experimental studies.

The approach is generally styled on the activities of a purposive or experimentally minded tutor, who experiments with various strategies and decision rules, presenting his items with specific goals in mind - this attitude enables appropriate constraints to be added. Several levels of strategies, student parameters and decision rules are envisaged, together with a well-documented information store, permitting the realization of a rich structure within which the system can be 'responsive' to student needs and peculiarities in a complex manner. The system may improve its teaching performance from experience (for example, from groups of students). The general line of approach envisages the handling of the multivariable probabilistic system by referring to a higher level goal, rule or measure, and allowing adjustment through the feedback of the parameters.

The problem is approached in the spirit that whatever can be devised in the first instance is bound to be modified in the light of further experience, consequently an attempt must be made to provide an overall structure, conceived with the expectation that this itself might need little modification, and the allowing of details, including organization of information, provision of strategies, decision rules and performance criteria, to be filled in by the experimenter, who may wish to test particular learning theories, or discover learning theories by experimentation. Consequently, two important complementary aspects in this system must be separately identified:

(a) The System Structure Itself - which allows processing of information to achieve adaptation (of the system to the learner), optimization, and allows the machine to learn to be a better 'teacher'.

(b) The Transformations between information to be imparted (in the student-machine interactions) and the student parameters.

Whereas (a) is reasonably explicit, (b) relates to an area in which much has yet to be discovered, and indeed it is to provide insight into these areas that has motivated the devising of (a). This circular relationship between (a) and (b) makes very apparent the need for both aspects to progress in an iterative manner - progress in one aspect leading to progress in the other, the net effect resulting in a general advance. The need for an experimental approach based on all relevant knowledge seems clear.

System Characteristics

The following are some of the characteristics (already discussed previously) which are desirable:

(i) The system must be responsive (i.e. adapt), to a whole range of student characteristics. Amongst other requirements, this necessitates an extensive student model structure, able to represent accurately many aspects of student behaviour.

(ii) A general structure which can handle a variety of topics or lessons, simply by allocating values to various 'item parameters', which then specify actual items.

(iii) A system which models the behaviour of the expert tutor, experimenting with various strategies (choosing favourable approaches), and continually searching for student difficulties which may arise. In this sense, the system is desired to play an active role in the teaching interaction.

(iv) The system should be amenable to external control, so allowing experimentation with a variety of strategies, decision rules, performance criteria, student parameters, and so on, thereby permitting continual optimization.

(v) There should be a structure which allows self-optimization, once certain desired 'teaching objectives' are specified.

(vi) The system should allow the learner to ask questions and seek information which he requires.

(vii) Appropriate principles of self-organization, whereby constraints and action capabilities are devised to ensure appropriate functioning, should be applied. Such a general system should be an aid in areas of Educational Research, and consequently should allow extensive monitoring of its operation and information flow.

5. CONCLUSIONS

It is our view that to try to formalize the output of a human teacher (for example by formalizing 'theories of learning') and then to employ such formalization in a mechanized teaching system - as essentially is being attempted in some work - is a practically fruitless approach. We consider that the 'machinery of teaching' should be studied, not the output of this 'machinery'. That is, we should try to discover those constraints and action capabilities which, when incorporated in a self-organizing teaching system, lend the system to a stable situation, interpreted as a goal, whereby a student learns because of the interactions with the system. The system should be capable of continually increasing its behaviour repertoire by learning.

The discussions and elementary work reported in this paper have been addressed to the problem of eventually realizing such a system.

6. REFERENCES

AMAREL, S., (1966), "On the Mechanization of Creative Processes", IEEE Spectrum, Vol. 3, No. 4, pp 112-114.

ASHBY, W.R., (1952a), "Design for a Brain", London: Chapman & Hall, Second Edition Revised, 1960.

ASHBY, W.R., (1952b), "Can a Mechanical Chess Player Outplay its Designer?" British Journal for the Philosophy of Science, Vol. 3, pp 44-57.

ASHBY, W.R., (1962), "Principles of the Self-Organizing System", in Von Foerster, H. & Zopf, Jr., G.W., Editors, "Principles of Self-Organization", Oxford: Pergamon Press, pp 255-278.

BEER, S., (1967), "Decision and Control", John Wiley: London and New York.

BITZER, D.L., HICKS, B.L., JOHNSON, R.L., and LYMAN, Elisabeth R., (1967), "The Plato System: Current Research and Developments", IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No.2, June, pp 64-70.

BRUNER, J.S., GOODNOW, J.J., and AUSTIN, G.A., (1956), "A Study of Thinking", New York: John Wiley.

CRUTCHFIELD, R.S., (1965), "Instructing the Individual in Creative Thinking" in "New Approaches to Individualizing Instruction", Educational Testing Service, Princeton.

GOLDBERG, A.L., TONDOW, M., and BUSHNELL, D.D., (1966), "The Computer in Education - Some Examples", Proc. IEEE, Vol.54, No.12, December, pp 1656-1662.

KANEFF, S., (1967), "Mechanization in Teaching-Learning Interactions", Journal of the Association for Programmed Instruction, Vol.2, No.2, October, pp 64-71.

MINSKY, M., (1961), "Steps Toward Artificial Intelligence", Proc. IRE, Vol.49, No.1, January, pp 8-30.

PASK, G., (1957), "Automatic Teaching Techniques", British Communications and Electronics, Vol.4, April, pp 210-211.

PASK, G., (1958a), "Teaching Machines", Proceedings of Second Conference of International Association of Cybernetics, Namur, pp 962-978.

PASK, G., (1958b), "Electronic Keyboard Teaching Machines", Education and Commerce, Vol.24, pp 16-26.

PASK, G., (1960a), "The Teaching Machine as a Control Mechanism", Transactions of the Society of Instrument Technology, Vol.12, No.2, June, pp 72-89.

PASK, G., (1960b), "Adaptive Teaching with Adaptive Machines", in Lumsdaine, A.A., and Glaser, R., "Teaching Machines and Programmed Learning: A Source Book", Washington D.C.: National Education Association, U.S., pp 349-366.

PASK, G., (1961), "An Approach to Cybernetics", London: Hutchinson.

PASK, G., (1962), "Interaction Between a Group of Subjects and an Adaptive Automaton to Produce a Self-Organizing System for Decision Making", in Yovits, M.C., Jacobi, G.T., and Goldstein, E.D., Editors, "Self-Organizing Systems - 1962", Washington D.C.: Spartan Press.

PASK, G., (1965), "Teaching as a Control-Engineering Process", Control, January, pp 6-11, February, pp 69-72, March, pp 153-157, April, pp 211-215.

PASK, G., (1967), "The Control of Learning in Small Subsystems of a Programmed Educational System", IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 2, June, pp 88-93.

PASK, G., and LEWIS, B.N., (1962), "An Adaptive Automaton for Teaching Small Groups", Perceptual Motor Skills, Vol. 14, pp 183-188.

PIAGET, J., (1957), "Logic and Psychology", New York: Basic Books.

RATH, G.J., (1967), "Computer Teaching - 1967", IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 2, June, p 59.

SOLOMONOFF, R.J., (1966), "Some Recent Work in Artificial Intelligence", Proc. IEEE, Vol. 54, No. 12, December, pp 1687-1697.

STOLUROW, L.M., (1961), "Teaching by Machine", Co-operative Research Monograph, No. 6, Washington D.C.: U.S. Department of Health, Education and Welfare.

VLADCOFF, A.N., (1968), "Application of Certain Principles of Self-Organization to Teaching System Structures", The Australian National University, Department of Engineering Physics Publication EP-T4, May.

VON FOERSTER, H., (1960), "On Self-Organizing Systems and Their Environments", in Yovits, M.C. and Cameron, S., Editors, "Self-Organizing Systems", New York: Pergamon Press, pp 31-50.

Publications by Department of Engineering Physics

No.	Author	Title	First Published	Re-issued
EP-RR 1	Hibbard, L. U.	Cementing Rotors for the Canberra Homopolar Generator	May, 1959	April, 1967
EP-RR 2	Carden, P. O.	Limitations of Rate of Rise of Pulse Current Imposed by Skin Effect in Rotors	Sept., 1962	April, 1967
EP-RR.3	Marshall, R. A.	The Design of Brushes for the Canberra Homopolar Generator	Jan., 1964	April, 1967
EP-RR 4	Marshall, R. A.	The Electrolytic Variable Resistance Test Load/Switch for the Canberra Homopolar Generator	May, 1964	April, 1967
EP-RR 5	Inall, E.K.	The Mark II Coupling and Rotor Centering Registers for the Canberra Homopolar Generator	Oct., 1964	April, 1967
EP-RR 6	Inall, E.K.	A Review of the Specifications and Design of the Mark II Oil Lubricated Thrust and Centering Bearings of the Canberra Homopolar Generator	Nov., 1964	April, 1967
EP-RR 7	Inall, E.K.	Proving Tests on the Canberra Homopolar Generator with the Two Rotors Connected in Series	Feb., 1966	April, 1967
EP-RR 8	Brady, T.W.	Notes on Speed Balance Controls on the Canberra Homopolar Generator	Mar., 1966	April, 1967
EP-RR 9	Inall, E.K.	Tests on the Canberra Homopolar Generator Arranged to Supply the 5 Megawatt Magnet	May, 1966	April, 1967

No.	Author	Title	First Published	Re-issued
EP-RR 10	Brady, T.W.	A Study of the Performance of the 1000 kW Motor Generator Set Supplying the Canberra Homopolar Generator Field	June, 1966	April, 1967
EP-RR 11	Macleod, I.D.G.	Instrumentation and Control of the Canberra Homopolar Generator by On-Line Computer	Oct., 1966	April, 1967
EP-RR 12	Carden, P.O.	Mechanical Stresses in an Infinitely Long Homogeneous Bitter Solenoid with Finite External Field	Jan., 1967	
EP-RR 13	Macleod, I.D.G.	A Survey of Isolation Amplifier Circuits	Feb., 1967	
EP-RR 14	Inall, E.K.	The Mark III Coupling for the Rotors of the Canberra Homopolar Generator	Feb., 1967	
EP-RR 15	Bydder, E.L. Liley, B.S.	On the Integration of "Boltzmann-Like" Collision Integrals	Mar., 1967	
EP-RR 16	Vance, C.F.	Simple Thyristor Circuits to Pulse-Fire Ignitrons for Capacitor Discharge	Mar., 1967	
EP-RR 17	Bydder, E.L.	On the Evaluation of Elastic and Inelastic Collision Frequencies for Hydrogenic-Like Plasmas	Sept., 1967	
EP-RR 18	Stebbens, A. Ward, H.	The Design of Brushes for the Homopolar Generator at The Australian National University	Mar., 1964	Sept., 1967

No.	Author	Title	First Published	Re-issued
EP-RR 19	Carden, P.O.	Features of the High Field Magnet Laboratory at the Australian National University, Canberra	Jan., 1967	
EP-RR 20	Kaneff, S. Vladcoff, A.N.	Self-Organizing teaching Systems	Dec., 1968	
EP-RR 21	Vance, C.F.	Microwave Power trans-mission Ratio: Its Use in Estimating Electron Density	Feb., 1969	
EP-RR 22	Smith, B.D.	An Investigation of Arcing in the Electrolytic Switch/Test Load Used with the Homopolar Generator	Oct., 1969	
EP-RR 23	Inall, E.K.	Use of the Homopolar Generator to Power Xenon Discharge Tubes and some Associated Switching Problems	Mar., 1969	
EP-RR 24	Carden, P.O.	Pivoted Hydrostatic Bearing Pads for the Canberra Homopolar Generator	Dec., 1969	
EP-RR 25	Carden, P.O. Whelan, R.E.	Instrumentation for the Australian National University 300 kilogauss Experimental Magnet	Dec., 1969	

Copies of this and other Publications (see list inside) of the Department of Engineering Physics may be obtained from:

The Australian National University Press,
P.O. Box 4, Canberra, A.C.T., 2600.
Australia.

Price: $A1.00

National Library of Australia Card Number and ISBN 0 85584 000 5