A General Formal Model for Representing Test Item Results.

David Barzilay

A thesis
in
The Department
of
Education

Presented in Partial Fulfilment of the requirements
for the Degree of Master of Art at
Concordia University
Montreal, Quebec, Canada

June 1996

0-612-25947-1

ABSTRACT

A General Formal Model for Representing Test Item Results.

David Barzilay

The measured states of knowledge, in terms of concepts and associated entities can be represented and expressed in a variety of ways: pictorial, diagrammatic, symbolic or linguistic. A symbolic approach within the framework of matrix algebra is proposed in this study as a good representation for modelling the learner's state of knowledge. This form of representation allows for mathematical manipulations and operations on the entities that make up these matrices. This permits the determination of knowledge gain from successive tests that are administered and recorded within this framework. The technique is very general and flexible.

ACKNOWLEDGMENTS.

I would like to thank Dr. Gary B. Boyd for having had the confidence and trust to allow me to pursue personal ideas relating to this thesis and his overall interest and guidance during the course of this work.

I would also like to thank Dr. P. D. Mitchell for numerous enlightened discussions during my period of study and his unfailing skill at encouraging students to achieve their goals.

Special thanks to my dear wife, Sandra, for her patience, tolerance and unbounded love without which this work would have never been completed.

TABLE OF CONTENTS.

Chapter 1: Literature Review

Introduction.

An issue of concern that often arises in school and colleges is how to track and monitor the progress of the learner or a group of learners. Difficulties arise when the number of concepts or items to be learned and the relationships that may exist between them become quite numerous. Hence, some kind of representation is required which will adequately display labels of the concepts to be learned and one where those deemed mastered on tests can be distinguished from those yet to be administered.

From an educational technology perspective of effectiveness and efficiency this representation must be constructed so that results from test items administered can be readily exhibited and used. Thes test items results will of course depend on the specific measures chosen.

The use of concept maps by learners and teachers has been investigated by numerous authors from both the conceptual and procedural view point, Ausubel (1968), in Educational Psychology, L.N. Moreira (1979). The fundamental idea is that an individual can form a mental map consisting of various related parts or components in order to facilitate understanding some body of knowledge; this map may be made at some conscious level in the form of some imagery as an aid in conceiving scientific ideas, (Miller, 1987), or it may be some kind of symbolic representation, or possibly it may be formed at a subconscious level. If such cognitive maps are formed then comprehension can be enhanced. This idea has been applied in teaching college level courses, and in facilitating learning processes, for example in the teaching of courses in biology, in High School, (Schmid and Telaro, 1990) and at the collegiate level, (Taylor, 1990).

Thus a student may usefully constructs some sort of cognitive map of some material that has been taught or of some material that he or she has read. Similarly a teacher may form some cognitive map with respect to some topic of interest or some material that is to be transferred to a student or a group of students.

Considering those teaching situations which involve a student-teacher interaction, it is quite possible that the cognitive maps formed by the two individuals may not be similar or may have structures that are in some conflict with one another. From the point of view of learning achievement this is an undesirable situation. An ideal learning process is one in which there is mutual understanding between the student and the teacher. By mutual understanding it is meant that both student and teacher agree in their ways of exhibiting the meaning of some topic of discourse or the procedures involved in acquiring that meaning.

The notion of concept maps can be formalized by using concept networks. With these structures it is possible to identify a set of concepts and then create a set of directed links which will give a measure of the relationship between them. This structure enables one to determine some various paths possible which will connect one set of concepts to others.

A possible mode of handling concept networks relies on expressing them in terms of concept matrices, and in view of the vagueness that may be associated with the elements and links representing the concepts, fuzzy logic may have to be applied, along with its operational rules, in order to obtain results of final outcomes. In computer assisted learning, for instance, a formal connection-matrix based on language and a fuzzy logic based lesson is advocated, typically as a means of implementing and exploiting learner models, Boyd and Mitchell (1992).

The use of concept matrices is here suggested for the external representation of a state of knowledge or understanding of a learner and in an analogous manner that of the teacher or instructor, in terms of the concepts that form a particular body of knowledge. A set of operational rules will have to be developed in order to prescribe and obtain final learning outcomes. A recourse to fuzzy logic, with some variations, will undoubtedly eventually be required. Within this scheme, the most relevant characteristics of the learner and teacher will also be incorporated.

Literature Review.

In dealing with knowledge representation and the operational procedures which are associated with them, much work has been done within the domain of education and cognitive psychology. However, many of the mathematical procedures used rest on probability notions. Thus, statements about learning outcomes as a result of instruction given are expressed in terms of probabilities.

However, a significant shift has been made in the 1960's in the educational field in particular in defining school curricula in terms of discrete operational objectives, (Lipson, 1967). Each of the objectives in these curricula amounts to a class of behavior which can be generated by a given set of rules, such as the ability to perform various arithmetic operations, as an example. Scandura (1976) advocated deterministic theorizing in structural learning. The first of his three partial theories deals with how to characterize knowledge. Operational objectives fall into that deterministic characterization of knowledge. A list type of characterization of this sort has the advantage that it requires but a simple performance mechanism to determine if it has been mastered.

Concept graphs, as defined by Brachman (1985), are formal objects used to represent objects, attributes and relationships of the domain being modeled. A concept graph exactly reflects these representations by having object classes as nodes and relationships as edges, Dumslaff and Ebert (1992). Graphical representation has been a widely used approach in order to describe the conceptual structure of some topic to be learned in the educational field. Graph representation structures are useful for their pictorial representation and often enhance the learning process in that they allow for a clearer perspective on how a topic is structured in terms of its conceptual components.

Greeno (1972) has proposed a pictorial scheme to describe the relationship among several related concepts, the fundamental process being that two or more concepts lead to another one. From the mathematical viewpoint, it is an application of graph theory. One of its uses is for the analysis of problems and problem solving. As an illustration he cites an example involving the concepts of :

l = length, w = width, h = height, v = volume, A = area,

m = mass, d = density, s = speed, a = acceleration, f = force

and the relationships u are given by :

u_1 corresponding to the formula $v = l \times w \times h$

u_2 corresponding to the formula $A = l \times w$

u_3 corresponding to the formula $v = A \times h$

u_4 corresponding to the formula $d = \frac{m}{v}$

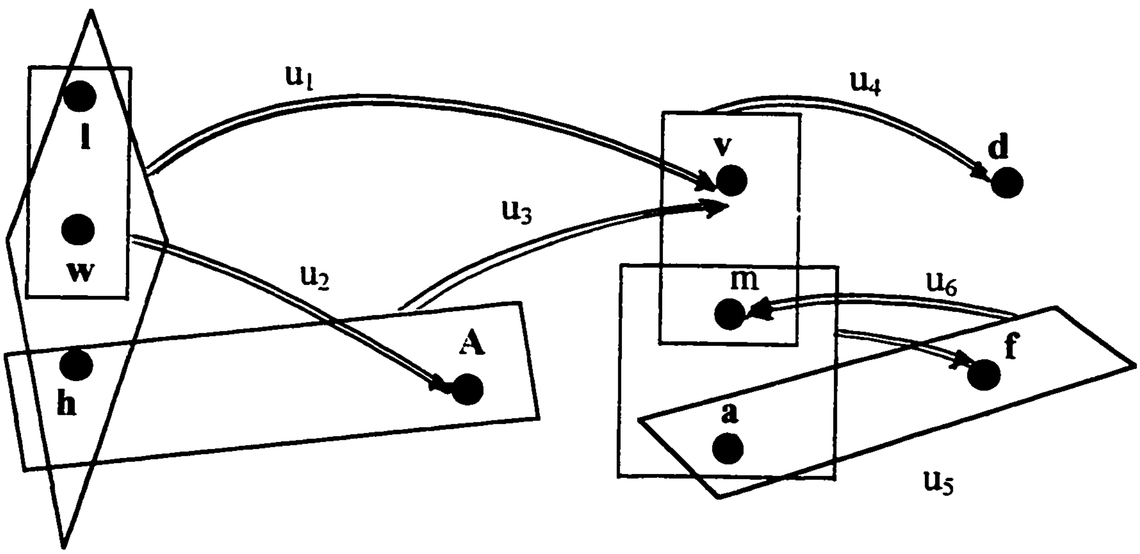
u_5 corresponding to the formula $f = ma$

u_6 corresponding to the formula $m = \frac{f}{a}$

The Greeno's pictorial representation appears as shown in the figure 1.1

Figure 1.1

Greeno's Pictorial Representation.



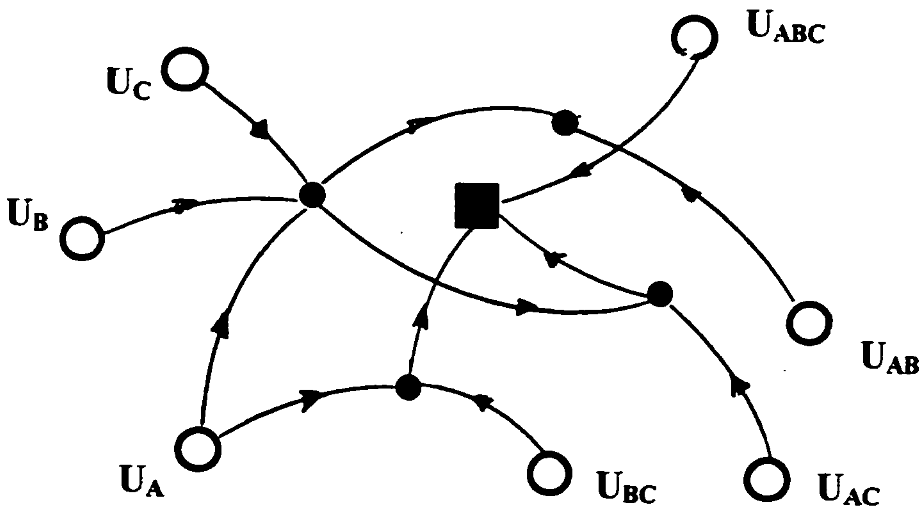
This graph theoretic approach is weak in that it does not describe the concepts themselves. It works mainly on the basis of connections that are known to exist among them. It is always possible in principle to define other connections and it would be a matter of further investigation to determine if these connections or relationships actually do occur. This approach has the advantage of flexibility. However, it can become extremely cumbersome when the number of concepts increases and as a consequence the combinations of their possible relationships increases exponentially.

Another graphical structural knowledge representation is based on work by Lamb (1966) and Reid (1968) who called this approach a "relation network". The basic idea is that concepts are connected by a "relation" symbolized at the interaction point. One may then construct a complex diagram showing all the interactions.

Another form of graphical knowledge representation was devised by Pask (1968), who named it the "entailment mesh" representation. A set of concepts and group of concepts form sets, U. These sets are shown as nodes. A problem is posed by X under a rule Ω and is solved by applying the correct operation, a string operations or a complex one. Correct operations are depicted as directed arcs which lead to a unique and central node, regarded as the solution. The description of the correct operation paths is symbolized as I(T). An example of this for three concepts A, B, C, is shown in the entailment mesh of figure 1.2.

Figure 1.2

Pask's Entailment Mesh Representation.



○ = Nodes- State of knowledge.

● = Other statements of knowledge.

——— = Correct operations

■ = Central final node

The above scheme is adequate for representing a knowledge structure of moderate size and how different combinations of concepts can lead to other states. Extensive labeling is required and the structure may be quite complex diagrammatically as the number of concepts increases. Also specific operations on each of the nodes must be specified and the directed arcs and paths which lead to the final solution. The directed arcs are a complex set of operations which must also be specified.

The transition from a graphical representation to a set theoretical one is quite a natural one. In fact many of the diagrammatic schemes used must be accompanied by some additional structure which explains and augments the structure. Symbols and alphabets are necessary to describe entities within that structure and possible relationships that may exist between them.

One begins by first constructing some cognitive map, which is expressed as directed graphs, with possible cycles feedback. The directed edges from one concept C_i to another C_j can indicate how much C_i causes, or the extent to which it is related to C_j. The time varying concept function $C_i(t)$ measures the non negative occurrence if some event, which may be fuzzy, Kosko (1986). Cognitive maps, which model the world as a collection of classes and causal connections between them have been applied in a variety of areas such as to model gastric appetite behavior, and popular political development, Taber (1987, 1991), also to analyze electrical circuits (Styblinski, 1988), and Chen (1988) in decision making situations. Kosko (1990) describes the operational process of going from the cognitive map to a matrix formulation in order to determine the effects that one entity would have on the entire structure and how that structure would evolve.

A fuzzified matrix approach has been found to be most useful in particular situations involving entities which have vague characteristics. The fuzzy logic operational rules are applied in the computational procedures. Such representations have been applied in document retrieval mechanisms by Her (1983) and query type problems by Kame (1990). These works can be related to a learning situation in so far as the process is concerned. That is to say, one begins with little or no knowledge and as the process evolves the desired state of learning is achieved.

Problem Description.

A body of knowledge which has to be imparted to a learner may and often contains many concepts and sub concepts associated with it . The problem that arises is how to describe this set of concepts or learning items in some systematic, compact and easily displayable and retrievable form, from which it is then possible operationalize it in some teaching situation and and provide a mechanism which testing can be performed so that it may be possible to monitor the learning outcomes as the learning process evolves.

The problem can be categorized in several parts,

a) How to express and build this knowledge structure within the framework of a matrix representation.
b) How to define and construct the relationship among the topics, expressed as elements of these concept matrices.
c) How to construct an operational mechanism which will involve the interaction of a set of matrices.
d) How to devise testing measures which are to be used when test items are administered, so that learning outcome can be evaluated.

Pask (1968), for instance, has devised an elaborate laboratory methodology for the learning processes, utilizing "entailment meshes". For a more formalizable description, a different approach is required. The approach must be such that the formalism must be expressed in such a manner that the operations between entities are clearly defined as well as the procedures to be followed. Adopting the view that " learning involves the construction of relationships among topics within a knowledge structure", (Mitchell and Grogono, 1993), this can then be formulated within the framework of the utilization of the a matrix representation.

One of the features that the proposed operational framework and representation should have is the possibility of including some student characteristics. This aspect is often ignored in the course of a conversation, be it a tutorial one or otherwise, or indeed in a conventional classroom teaching situation. This is due in part to the fact that teachers cannot possibly cope with the diversity of types of students within a classroom, and also due to the fact that there is insufficient domain expertise.

<u>The method</u>.

The notion of concept maps will be used as a point of departure. These being diagrams indicating concepts and their interrelationships by node labels and edges. The resulting maps can represent a person's conceptual organization of some given topic. However Pask's formalism, in terms of his conversation theory, and operationalized by his entailment meshes, will be used only as providing some very basic ideas in order to develop the required knowledge structure in terms of concept networks, expressed by the matrices.

The proposed structural scheme will be one which resembles, in some very limited aspects, an Intelligent Tutoring System (ITS), and one of its features will be such as to be able to incorporate some of the characteristics of the learner and instructional procedure. In this respect the proposed structure will have to be sufficiently flexible in order to adapt itself to a particular student. Thus the structure would have to be such as to represent, as accurately as possible, the version of the subject matter in the student's concept map, (Mitchell and Grogono, 1993).

The student's concept map will largely be a subset of the teacher/instructor's concept map, for a one directional tutorial conversation. i.e. one in which the student acquires all of his or her knowledge from the teacher. The requirement imposed here is that a minimal difference results between these two maps after learning has taken place.

Some Basic ideas of Pask's Conversation theory.

In Pask's conversation theory a conversational domain can be constructed with respect to any arbitrary subject matter. One considers two participants say A, the student, and B, the teacher. In this case the conversation is said to be a tutorial one. This in the sense that they both learn as a result of conversing about the domain, the domain being something to be discovered or learned about.

The conversation takes place in a potentially formalizable language L which is stratified into at least two levels of discourse: $L = L^{o}, L^{*}$
During the conversational act, commands and questions may be assumed at L^{o} or at L^{*} levels.
At the L^{o} level of discourse, statements are of the form, "I am doing this" and commands are of the form, "do something ", or "solve a problem". Questions are of the form, "give an explanation of...."

At the L^* level of discourse, commands are of the form, "learn to solve a problem" or "construct a process that solves this problem". Statements would be of the form : "I am constructing a ..." and questions demand the "explanation of how the process was constructed".

In general it is taken that that each A and B have at their disposal a repertoire of procedures which they can employ. Thus if a topic T_i is under discussion or investigation, then, in Pask's notation,

$Proc_A{}^o (T_i)$ represents - a procedure in A's repertoire that explains topic T_i

$Proc_A{}^* (T_i)$ represents - a procedure in A's repertoire that explains how $Proc^o (T_i)$ is learned, [and if applied to $Proc^* (T_i)$ itself may be regarded as A's reconstruction or re-explanation of A's memory of topic T_i .

If there is "agreement" between A and B, this is represented as:

$Proc_A{}^o (T_i) \Leftrightarrow Proc_B{}^o (T_i)$ similarly $Proc_A{}^* (T_i) \Leftrightarrow Proc_B{}^* (T_i)$

where T is the domain in which $Proc_A$ and $Proc_B$ are able to operate.

Thus if there is a doubled layer agreement, then it is said that there is "understanding" of topic T in L . i.e. this is Pask's definition of understanding.

It is taken that A and B can reproduce or reconstruct Con (T) from a set of components or subconcepts, labeled, a_i and b_i , $i = 1,2,3,\ldots\ldots. p$ respectively.

It is posited that optimal conditions under which understanding takes place is when matching occurs. That is to say, a given learning strategy must match a given teaching strategy. Thus, if a mismatched teaching strategy is employed, i.e. one that belongs to a class that's distinct from the class of learning strategy that a student would adopt, then virtually no relevant understanding takes place.

As stated earlier, in Pask's conversation notions, if $Proc_A^*(T_i) \Leftrightarrow Proc_B^*(T_i)$ then understanding has occurred. In essence this means that if B communicates to A his procedural derivations of constructing T_i and A shows agreement, then B's information, when processed in A's mind, yields the same result that A would have obtained with his own procedural construct of $Proc^*(T_i)$. "Understanding", from the operational sense, means matching A's procedural repertoire with B's at the the two levels, L^o and L^*.

Aspects of the proposed representation.

In the context of the work proposed for this thesis, with respect to the level of discourse, as expressed by Pask, those being L^0 (Lev 0) and L^*(Lev 1), only level L^0 will be considered as an application example. However the proposed structure and its subsequent operative mode will be sufficiently general that any two level of discourse in almost any formal language could be applied to it.

The language of discourse will depend to a large extent on the teaching method, styles, or instructional designs chosen, that is, the particular conversational mode. Since the learning outcomes have characteristics which cannot always be precisely observed, then measured fuzzy variables and possibly fuzzy logic operational rules will have to be invoked.

The underlying element in the structure and use of fuzzy logic is the "membership" function, $\mu(x)$, relating to some fuzzy set. Thus the approach and procedure utilized in determining this function must be taken with utmost care and rigor. The two variables involved are on the one hand the degree of membership to which a given value in the fuzzy subset is taken to belong, expressed on the unit line interval [0,1], and on the other hand the set of values within the fuzzy subset.

In the context of this thesis, the outcome of the learning process could possibly form a fuzzy set and the collection of hedges such as: very little, little, fair, good, average, very good, excellent, associated with this set would form the fuzzy subset. It seems that one cannot escape this approach since one must admit different levels of understanding, certainly during the learning process itself.

In principle, if one can quantify the procedures Proc $^{*}(T_i)$, for A and B, then one could consider the difference, D, between the two procedure constructs,

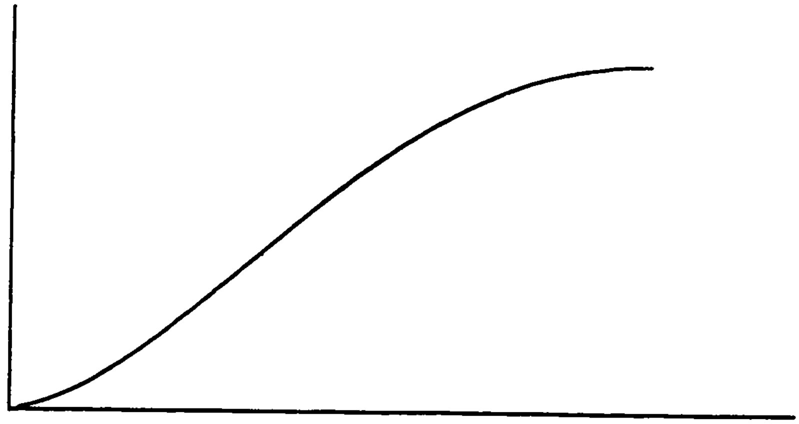
$$D = \text{proc}^{*}_{A} - \text{proc}^{*}_{B} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 1.1$$

and under suitable normalization conditions one can define, D = 0, implying full understanding, and D = 1 implying total absence of understanding.

One can speculate as to what function the learning outcome, expressed as U, can be adequately represented by. This might take the form of some well established learning curve. It is proposed here that, this learning curve assume the shape of a logistic curve, shown in figure 1.3.

Figure 1.3

Learning logistic curve.



The above curve can be described algebraically by an equation of the type,

$$U = \frac{1}{\sqrt{c_1 + \frac{D^2}{c_2}}} \qquad \text{.......................... 1.2}$$

with D as defined by equation 1.1, and c_1, c_2 are constants relating to the contextual situation. The numerical values of these constants can be obtained during the curve fitting procedure from a set of experimental data collected. One will note that the learning curves, of the type shown in figure 1.1, closely resembles curves obtained for fuzzy variables such as linguistic terms and hedges. An example would be of the form, " the degree of understanding of " a concept Con (T).

Equation 1.2 is but one of a variety of mathematical expressions that could be utilized to describe such curves such as the one shown in figure 1.1. The characteristics of an equation of this type are that it is simple to use and that the constants c_1, c_2 are adjustable to a given contextual situation. This would be necessary in view of the fact that learning curves may depend on the learning topics.

The parameter D, as defined by equation 1.1, is consistent with the notion that learning is relative to some defined body of knowledge. Thus, comparisons with what is regarded as the "knowledgeable source ", say $Proc_B$, or to what one might define as the absolute state of knowledge are necessary. The learning outcome variable, U, from testing can be described on the line interval [0,1], or as a percentage value.

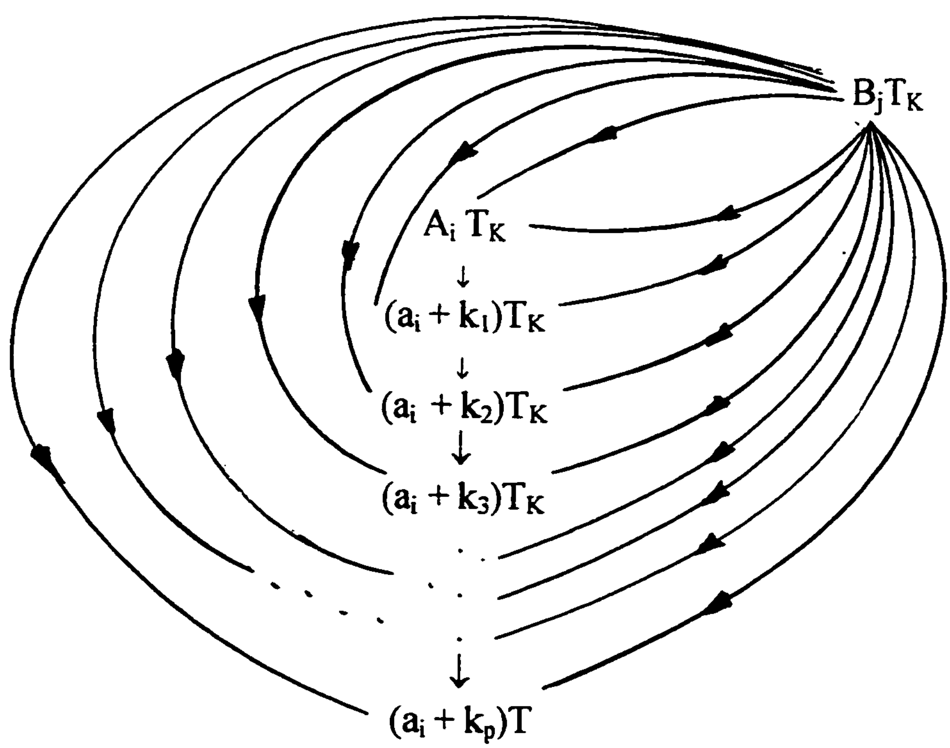
A Diagrammatic Representation.

A concept acquisition process can be described in a diagrammatic manner. In the simplest case one can consider a one way transmission process in which only one of the participants constructs "knowledge" and transmits it to another participant who acquires understanding. In this situation B possesses the required knowledge (or constructs it) which is imparted to A. In this sense B's knowledge of some topic T_i, expressed as $Proc_B^*(T_i)$ becomes the point of reference with A's $Proc_A^*(T_i)$.

This tutorial conversation, or one way instructional discourse, is described in figure 2.1

Figure 2.1

One way Instructional Discourse.



This leads to the equivalent relation,

$$\text{Proc}_A^*(T_i) \Leftrightarrow \text{Proc}_B^*(T_i)$$

In the above diagrammatic description, the curved lines represent some particular process or procedure for which knowledge is imparted or transferred from person B to person A. The vertical downward arrow represents the higher level of understanding that A acquires, as a result of this conversation with B. The curved lines, drawn from A to B, can represent communication or again some process or procedure that the new level or state acknowledgment has been achieved.

In the flow chart shown, the process may be "stuck" at a given level and the cycle repeating itself until the higher state of knowledge or understanding has been attained. It is assumed in this representation that B possesses all the "knowledge" that A will acquire during this tutorial conversation process.

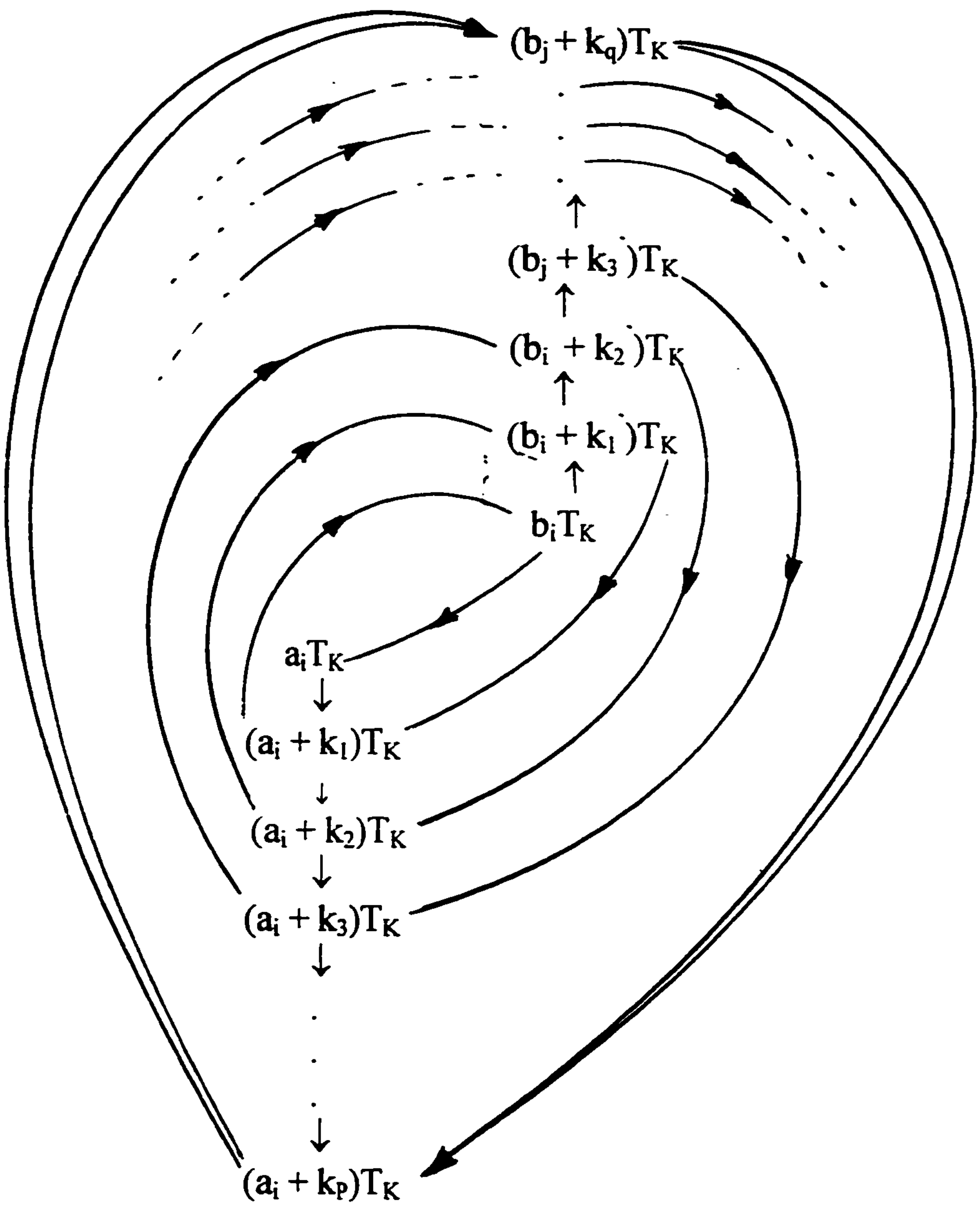
The coefficient A_i can be made to represent the degree of understanding, that person A has for a specific element or component associated with Con (T_i). In this learning process the objective is to increase that degree of understanding, or, decrease the fuzzy value of A's understanding of topic T_K. The k_i coefficients, i = 1,2,3, ... p, represent some incremental increase in understanding of that component of the topic. These incremental increases need not necessarily be of equal values, and in practice, in most cases will not be.

The above scheme can be extended to the situation where both A and B learn from each other. Thus it is taken that A and B, individually, possess some of the required knowledge, or each have partial understanding of some concept Con(T). The initial starting point may be with either A or B.

This two way instructional discourse is described in figure 2.2.

Figure 2.2

Two way Instructional Discourse.



The curved lines that connect the levels of knowledge reached during the instructional discourse could be drawn in different graphic styles which

would represent different teaching processes such as conventional classroom setting, individualized learning or computer assisted learning. In the above diagram it is not suggested that the increments of states of knowledge or understanding reached during the discourse are equivalent at every level.
The above geometric representation assumes a simple kind of learning evolution, in its most general form, between two individuals A and B.

One can devise a geometric representations which would describe a specific learning activities and some specific learning instructions at different levels and stages. This would require a much more elaborate and complicated geometric representation, which would be quite feasible in principle but perhaps not quite so simple to apply.

An example of a knowledge acquisition process as represented by the geometric representations of figures 2.1 and 2.2 in the context of an instructional discourse, or a tutorial conversation between A and B would be in a situation where a conventional computer aided learning, CAL, is utilized. Thus, in figure 2.1 the computer is represented by B_jT_{BK} and the student by A_iT_{AK}.

Instruction takes place whereby B instructs A, and upon completing the required learning tasks A reports the results back to B, which then issues the next level of instructions, on determining from the report whether or not the given level of knowledge has been achieved. The process continues until agreement of understanding has been reached or the final state of knowledge has been attained.

Proposed Concept Matrix.

To account for the state of the learner during the knowledge acquisition process, and as a basis to ensure that optimum understanding is achieved, it is necessary to define the matrix representing the learner's, initial knowledge state. Here the general coefficients a_{ij} represent the initial state of knowledge of the learner with respect to the set of procedures that must be known in order to achieve Pask's condition of $Proc_A^*(T_i)$ Since it is assumed that the learner has no prior knowledge of the topic T_i, or perhaps very little, then these coefficients, a_{ij}, can be regarded as fuzzy subsets of the fuzzy set represented by this matrix. If person B's knowledge is also not complete then the corresponding matrix will also be a fuzzy set.

Since knowledge acquisition is a result of interaction between the two individuals, in an educational context, mathematically this means that one must consider the result of some logical operation between the two corresponding matrices, labeled M_A and M_B, which effectively represent the individuals. Thus a function φ maps a domain in B to a range in A. This association, which may be fuzzy, (M_A, M_B) represents a system structure, or a fuzzy system S. One can write this as,

$$S : I^Q \rightarrow I^P \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad 2.1.$$

where knowledge is represented by a Q dimensional hypercube, I^Q, containing all the crisp or fuzzy subsets of the domain space, (that of the teacher or instructor), or input in the universe of discourse, and I^P, the P dimensional hypercube containing all the crisp or fuzzy subsets of the range of the learner, (Kosko, 1993). Equation 2.1 suggests that one adopts the view that matrices M_A and M_B be regarded as fuzzy subsets, since the crisp sets are but a special case of the fuzzy sets, hence one can always incorporate this special case if the situation or context demands it.

Regardless of which operational definition one wishes to use for these matrices, the common feature that they possess is that their elements must necessarily be fuzzy or possibly be, this in view of the fact that one initially begins with some incomplete knowledge of the topic under investigation. Thus each of the coefficients can only partially describe the state of knowledge of a topic.

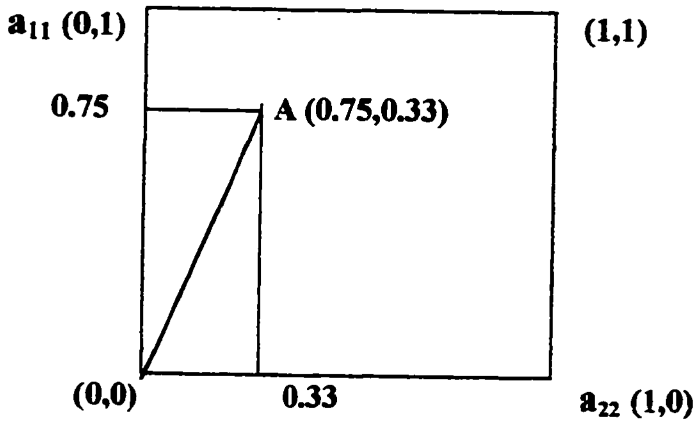
The coefficients described in these matrices have two fundamental functions. One, is to identify or label the items or concepts pertaining to a domain of knowledge. Two, to quantify with respect to some measure the level or state of knowledge or understanding reached. With respect to the hypercube, they become their dimensions. These shall be referred to as the primary dimensions of these matrices, or the fundamental concept of some topic, and the a_{ij} ($i \neq j$) coefficients as the secondary dimensions, the subconcepts. One can use a geometric representation, as used by Kosko (1993) to describe this.

As an example consider the simple case where the requirement of understanding consists of only two coefficients or the description for exhibiting a learning process requires but two descriptive components, say a_{11} and a_{22} with no overlap between them. If some vagueness exists in their knowledge or understanding, each of these can represented by a one dimensional axis with a range [0,1]. Thus a fuzzy subset A_P is a point in this two dimensional cube with respective coordinates, or fit values i.e. the degree of membership in each of the coefficients. Considering a numerical example,

$a_{11} = 0.75$ and $a_{22} = 0.33$. This can be represented geometrically by the diagram shown in figure 2.5.

Figure 2.3

2-D Fuzzy Hypercube.



The above is a geometric display showing the degree of understanding with respect to two of the primary components. One could convert the numeric range to percentages whereby a_{11} is understood to a degree of 75% and a_{22} to a degree of 33%. Full understanding would be represented by the vertex (1,1). One will note that the dimensions of this hypercube, which are representatives of the a_{ii} coefficients, does not include any overlap that may exist between these coefficients, or in fuzzy logic terms, the intersection of their membership values. In principle there is no reason why one cannot include any of the a_{ij} $(i \neq j)$ coefficients using the above geometric representation. It is a matter of simply redefining the dimensions of this hypercube.

In the proposed structure representing the concepts of some person A, and similarly that of a person B a set of elements, or coefficients, a_{ii} , b_{ii} , will represent the main concepts of some body of knowledge. This notation replaces the labels C_i described earlier and will be represented by a column matrices, written as.

$$M_{a_{ii}} = \begin{pmatrix} a_{11} \\ a_{22} \\ a_{33} \\ a_{44} \\ . \\ . \\ a_{pp} \end{pmatrix} \quad and \quad M_{b_{ii}} = \begin{pmatrix} b_{11} \\ b_{22} \\ b_{33} \\ b_{44} \\ . \\ . \\ b_{qq} \end{pmatrix} \qquad \ldots\ldots\ldots\ldots 2.2$$

An additional matrix M_A can be constructed in which an additional set of coefficients will represent set of subconcepts or a set of procedures required for the understanding of the primary concept. The membership functions or degree of association that these subconcepts have to the primary set of concepts can replace the directed links that appear in the concept network representation. Keeping in mind that these apply to a particular property of the association, since there may exist several properties of associations.

The matrix in this scheme, as mentioned previously, also expresses the knowledge state of the student. This will be the case if one regards the coefficients defined in this matrix as representing the state of understanding that the student has reached for a set of concepts or subconcepts within the defined domain. The matrix can in this restricted sense be used to model the student, where modeling of the learner is restricted to the definition related to the state of understanding acquired.

This representation can be compared, to some extent, with recent developments in student modeling utilizing Knowledge Space Theory (KST) of Falmagne and Deignon (1990). In Knowledge Space Theory, an area of expertise can be quantized into "items". These items may in fact be represented by the concept matrix coefficients. Thus an item may represent a task in which a student must perform, if the goal is to assess procedural knowledge, similar to Pask's $Proc_A^*$ (T). The quantization aspect only occurs in the setting up of the concept matrix, that is as defined in terms of its coefficients. The body of knowledge is thus characterized by the set of items (coefficients) called the domain. The student's knowledge state is defined as the collection of these items (coefficients), if they represent some procedural task, or a cognitive task.

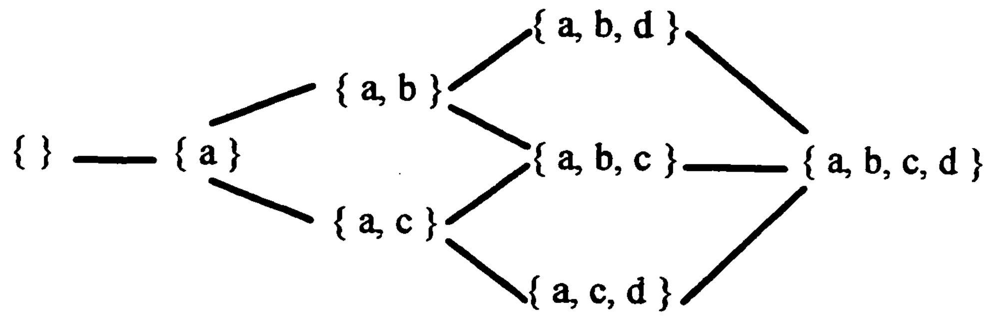
One of the characteristics of KST is that some inference is made when constructing the *knowledge structure*. In the example used by Villano (1992), suppose there are four items a, b, c, and d to be learned, where,

a. 4 x 7 = ?   b. 1/4 x 1/7 = ?   c. 0.4 x 7 = ?   d. 40% of 7 = ?

Mathematically all possible combinations of these items would form the knowledge space. However, due to the association between these items it might be inferred that one cannot understand item d without having an understanding of item a. Thus any state containing item d would also have to include item a. Prerequisite relations play a critical role in structuring this knowledge space. An example of the knowledge structure is shown in figure 2.6.

Figure 2.4

Example of Knowledge Structure.



In the context of the concept matrix representation this would be in the manner in which the matrix is evaluated as a whole. Thus the proposed matrix representation is an analogous manner is a construct yielding a knowledge space. Although a comparative analysis with Flamagnon et al knowledge spaces will not be made here, significant differences do exist between the two in terms of the constraints imposed on the structure via the prerequisite relations and direct relationship between the items.

The concept matrices representing the state of knowledge for persons A and B can thus be written in terms of all the coefficients as,

$$M_A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . & a_{1r} \\ a_{22} & a_{21} & a_{23} & . & . & a_{2r} \\ a_{33} & a_{31} & a_{32} & . & . & a_{3r} \\ a_{44} & a_{41} & a_{42} & . & . & a_{4r} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{pp} & . & . & . & . & a_{pr} \end{pmatrix} \quad M_B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & . & . & b_{1s} \\ b_{22} & b_{21} & b_{23} & . & . & b_{2s} \\ b_{33} & b_{31} & b_{32} & . & . & b_{3s} \\ b_{44} & b_{41} & b_{42} & . & . & b_{4s} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ b_{qq} & b_{q1} & b_{q2} & . & . & b_{qs} \end{pmatrix} \quad \ldots\ldots 2.$$

In the notation described in the above concept matrix, the coefficients a_{ii}, (a_{11}, a_{22}, a_{33}, . . .), represent the basic fundamental concepts of some body of knowledge. The remaining coefficients, a_{ij} , b_{ij} ($i \neq j$), along a given row represent the subconcepts associated with the fundamental one.

With this proposed structure one is essentially simplifying the concept network in that it does not take into account the overlap, or the subsets that may exist. The proposed structure does not exclude the possibility of expressing relationships that may exist between the coefficients. This aspect will be dealt with in a later chapter. The manner in which the matrices are to be evaluated will be of critical importance.

Chapter 3 : Mathematical Constructs.

Basic notions.

It is possible to define the coefficients as representing measurements of the state of knowledge that one has with regards to a set of specified concepts and subconcepts, as was mentioned earlier. In this context the state of knowledge can be determined by evaluating these matrices. If applied in this sense, then the manner in which the matrices are to be evaluated will be of critical importance. There are several methods that one can consider.

Given the fuzzy nature of the elements of these matrices, one common technique is the application of measure theory, Klir and Folger (1988). In principle one is at liberty to define any measure of fuzziness one wishes with respect to some defined criteria. Of the variety that have been used, and the most common one, is the one based on metric distances. For a particular case of a Euclidean distance, one has,

$$f(A)=\left(\sum\left|\mu_A(a_{ij})-\mu_C(a_{ij})\right|^2\right)^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots 3.1$$

More general metric distances can be adopted, notably the Minkowski class of distances given by,

$$f_W(A)=\left(\sum\left|\mu_A(a_{ij})-\mu_C(a_{ij})\right|^W\right)^{\frac{1}{W}} \ldots\ldots\ldots\ldots 3.2$$

where $\mu_A(a_{ij})$ is the membership value for that coefficient and $\mu_C(a_{ij})$ is its corresponding crisp set and $w \in [1,\infty]$. The value of w can be determined from criteria set for the definitions of sharpness and maximum fuzziness, (Klir and Folger, 1988). Also its value will depend on the specific geometric framework one wishes to deal with. As to which of these geometric representation is the most suitable would be a matter requiring further investigation.

Utilizing these notions, the "index of fuzziness" can thus be readily translatable to an "index of learning gain". One can note that in the case where person A learns from person B, then the number of coefficients in matrix M_A could possibly be equal but not greater than the number of coefficients in matrix M_B. It is possible however that the number of coefficients of matrix M_A could exceed those of matrix M_B in the particular case where the learner has many misconceptions about a topic and therefore many of these have to be corrected. This is often referred to as the " Bugs" model.

Once the concept matrices have been constructed, for the student/learner and the teacher/instructor, the M_A and M_B matrices, it would be necessary to determine how these two matrices are to interact. This means that one must determine the algebraic operational rules that govern the interaction of these coefficients, which will be some form of multiplication.

The matrix multiplication to be considered will be written as $M_A \otimes M_B$. It should be noted that the rules of multiplication will not follow the standard matrix multiplication rules. Since one will be dealing with quantities that represent vague notions of understanding, or incomplete ones, hence have fuzzy characteristics. Then it would seem that some variation of fuzzy multiplication rules will have to be adopted. These rules will be specifically defined later and will have to be such as to be internally consistent with the representation described.

The fundamental fuzzy multiplication rules are based on Maximum-Minimum conditions, (Terano, 1992), with respect to the elements in the fuzzy subsets. This however cannot hold in this context of this representation in view of the fact that one cannot ensure that maximum learning has occurred during a learning process or whenever a tutorial conversational exchange has taken place. In fact figures 2.1 and 2.2 clearly show that to achieve maximum learning or understanding, several intermediary stages of learning take place. These intermediary stages are necessary for one may have to take into account the occurrence or forgetting or having misunderstood.

The objective here is that one wishes to obtain a quantifiable measure of the learning that has occurred as a result of the interaction that took place between the student and teacher.

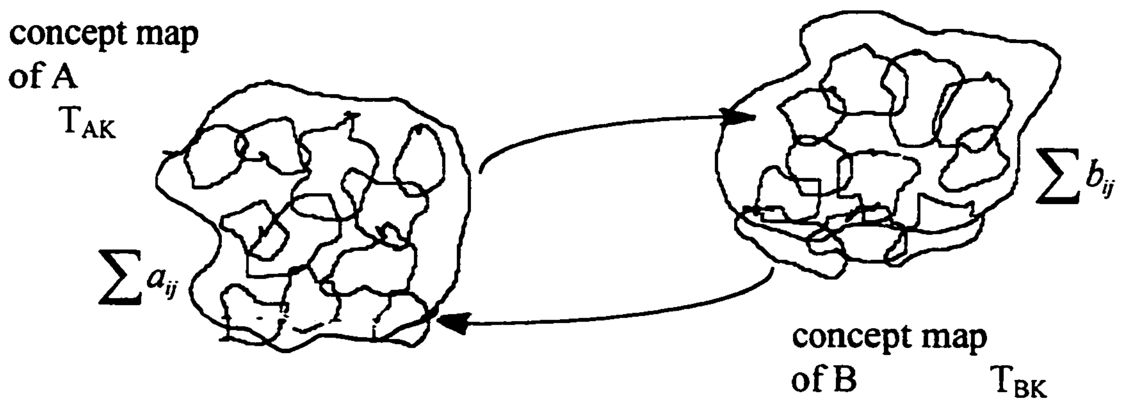
Description of the Coefficients.

In the hierarchical learning stages that were described in figures 2.1 and 2.2, the a_i and b_j coefficients represented a factor of the knowledge that exist in A's and B's minds relating to the same topic of discourse T_i, or a fuzzy value of that topic. However one should extend this notion by stating that the full knowledge of a given topic involves a complete set of sub concepts. Thus the index i may extend to several values, say i = 1,2,3,.......p. That is, the topic T_K is described in terms of a set of a_i coefficients, a_1, a_2, a_3, ...,a_p .These coefficients should however be regarded as fuzzy subsets of the set T_K. Furthermore one must consider the possibility that there may exist some overlap between these fuzzy subsets.

Using these coefficients, theoretically one can describe the concept that person A has in his mind of some topic T_K to represent the cognitive map formed in his mind, these can then in turn be expressed in a matrix form. Pictorially one can describe the interacting concept maps as shown in figure 3.1 With this description it is possible to consider topological properties of the elements that form the concept maps and the mapping that exists between them.

Figure 3.1

Interacting Concept Maps



One writes the set of coefficients a_{ij} and b_{ij} in a columns and rows and expressed in matrix form, as expressed by equation 2.3 in the preceding chapter. One now postulates that for a concept or topic T_K to be "understood", in the sense that the procedures $Proc^*(T_K)$ exist, then all the coefficients in the above matrices must exist and the a_{ij}'s must correspond to the b_{ij}'s. If initially some of the coefficients do not exist, at least in one of the persons' minds, then through appropriate interaction during the learning process between A and B, the necessary coefficients will come into existence as understanding is acquired. Thus interaction between the a_{ip} and

b_{iq} coefficients will give rise to other coefficients, or will result in an augmented numeric value from its preceding one. This situation was depicted in figures 2.1 and 2.2, shown earlier.

The notion of creating learning outcome coefficients is not a strange one in light of the fact that, initially, one may not be in possession of any particular knowledge of a given topic. The creation of these new coefficients are as a result of the conversation that is taking place.

In the operational context this means that some teaching operator, represented by the b_{ij} coefficient, from person B, operates on the knowledge state of person A, thereby increasing that person's state of knowledge by either augmenting its existent state, a_{ij} , or creating a new one if one did not exist at all. As was described in figure 2.1. The requirement is that one reaches the state whereby all the a_{ij} coefficients attain the value of $(a_{ij})_{max}$.

Although the matrices described by equation 2.3 appear as P x P and Q x Q in dimension, this need not necessarily be the case. What this means is that the procedure required for the understanding of a specific concept need not necessarily require the same number of procedural steps, or be composed of the same number of subconcepts, as that of another concept. In this context, the term procedures, in Pask's sense, is interchangeable with the term subconcept. Hence this matrix may have many zero elements in it. In Pask's terms, the set of a_{ij} coefficients represent the procedure in A's mind that enable him/her to construct or reconstruct the concept associated with T_K. The number of subconcepts required can be defined a priori or created as the process of learning is evolving.

One can symbolize the resulting learning outcome by the letter U, and express it as an algebraic expression given by.

$$U_A = \sum_i \sum_j a_{ij} = \text{Proc}_A^*(T_i) \quad \ldots\ldots\ldots\ldots 3.3$$

How well one can operationalize this matrix depends to a large extent on the topic to be learned and externalizability factors. It is quite evident that the more well defined the topic is, i.e one whose fuzziness is smaller, the more readily it would be applicable. Typical disciplines would be in the realm of the physical sciences and mathematics, also any learning tasks that only require the acquisition of some well defined skills, physical or cognitive.

The above description for the knowledge acquisition process, as formulated, makes the requirement that all the procedural acts described in the matrices A_i and B_j must exist. Whether one needs to actualize these procedures every time is a matter of practical requirements in a given context. For example, in the case of a mathematical procedure of integrating some function, one need not necessarily be required to go through all the intermediate steps in details in order to reach the final result. However it is understood that one has in ones possession the necessary skills to go through these procedures if called upon.

Augmented Coefficients.

It is necessary to explain the relationship between the representation as described in figures 2.1 and 2.2 with that of the coefficients in the knowledge acquisition matrices. During the course of a conversation between A and B, say a tutorial, both A and B can augment their knowledge, or understanding of some topic or subtopic T_i by some incremental amount k_j for A and $k_j^{'}$ for B.

These k coefficients may be well defined, discrete quantities, such as the steps required to learn the operation of a machine, or the steps required to solve a very well defined problem in mathematics or in the physical sciences or engineering.

In general, however, one can assign the coefficients by a fuzzy value in the interval [0,1] as one would in the social sciences and liberal arts where concepts are not always and easily described as crisp entities. Thus if the values of the coefficients in the matrices of equation 2.2, at a time t are written as a_{ij}^n and b_{ij}^n, then in the process of learning their values are augmented by amounts k_j and $k_j^{'}$ to new values a_{ij}^{n+k} and b_{ij}^{n+k} respectively. This of course depending on who is acquiring the knowledge. This can be written in an equation form, for matrix M_A, as

$$\begin{aligned} a_{ij}^n + k_j &= a_{ij}^{n+k} \\ b_{ij}^n + k_j^{'} &= b_{ij}^{n+k^{'}} \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots\ldots 3.4$$

If sufficient time is permitted for the learning process to take place, then one has to assume that these coefficients will each reach their maximum possible attainable values. i.e. $a_{ij}^{n+k_j} = (a_{ij})_{\max}$ and $b_{ij}^{k^{'}} = (b_{ij})_{\max}$

Properties and Characteristics of the Coefficients.

The procedural rules associated with the coefficients of this matrix representation will depend on the properties and characteristics of these coefficients. In this respect it is important to define these in an unambiguous manner. These are defined by the following :

1. a_{ii} This refers to the main elements or main concepts relating to a given topic to be learned or a body of knowledge to be understood.

2. $a_{ij} = a_i \cap a_j$ This reflects the intersection property between coefficients. It could refers to a sub component of the main component a_{ii}. It would be a subset of it. This typically would be the case if the main concept is defined entirely in terms of other concepts.
   More generally it could refer to a relationship that may exist between the coefficients. This would then give rise to a set of relational coefficients.

3. $a_{ij} \neq a_{ji}$ This is the noncommutative property of the coefficients. This arises from the labeling of the coefficients, where the subscript i refers to the main components.

4. $a_{ii} \geq a_{ij}$ This follows from the definition of the a_{ij} coefficients in property 2.

5. $a_{ii} \leq 1$ and $a_{ij} \leq 1$ This is due do the possible fuzzy character of the coefficients which are defined on the line interval [0,1].

6. if $a_{ip} = 0$ then $a_{ip}b_{iq} \geq 0$ This signifies that A can learn in the absence of any initial topic specific knowledge, a knowledge state of zero. If the resulting product is zero, it signifies that no learning has occurred.

7. $a_{ip}b_{ik} \neq b_{ik}a_{ip}$ This reflects the non commutative property of these coefficients. The learning outcome as a result of A learning from B is not necessarily equivalent the learning outcome of B learning from A.

8. $a_{ip}\, b_{ik} \geq a_{ip}$, if $b_{ik} \geq a_{ij}$. Since A learns from B then by definition the coefficients in matrix M_A must be less than those of matrix M_B, and its coefficients must by definition be augmented, by some increment k_i. if learning has occurred.

9. $\sum(a_{ip})_{coeff} \leq \sum(b_{iq})_{ceff}$ This refers to the number of elements in each of the matrices. Since person A has less knowledge than person B, then by definition matrix A must have less elements in its matrix. Conceivably many of the elements in it may have zero values in the initial stage, t = 0. The converse also holds if B has less knowledge than A.

10. If property 1. does not hold, then neither does property 2., in which case the learning matrix M_A can assume all the coefficients and the maximum possible number of elements in that matrix will be simply pq,

or $\sum_{i=p}^{i=1}\sum_{j=q}^{j=1}(a_{ij})_{coeff}$.

This means that there are no primary or main concepts, or learning items, defined within the described body of knowledge. Hence all the coefficients are independent of one another.

11. The learning matrix are defined in terms of primary and secondary dimensions. The primary dimension, D_P, given by the set of elements

$D_P = \sum_{i=p}^{i=1}(a_{ii})_{coeff}$ in the matrix. The secondary dimension, D_S, is given

by the set of elements $D_S = \sum_i \sum_j (a_{ij})_{coeff}$, $i \neq j$ in that matrix.

This definition is imposed in view of the fact that the a_{ij} $(i \neq j)$ coefficients are possible subsets or subconcepts of the a_{ii} coefficients as defined in property 2. Furthermore this definition of the primary and secondary elements allow for the description of a topic or concept Con (T) to be described in some hierarchical or systemic manner, if one chooses. The global dimension is defined as :

$$D = D_P + D_S = \sum_i \sum_j (a_{ij})_{coeff}$$

12. $l_{ij} = \sum \sum a_{ij} b_{ij}$ This defines the learning of a given component as the summation of the products of the coefficients a_{ij} and b_{ij}. The summation is that of ordinary arithmetic, however, since each of the coefficients could be expressed as fuzzy values in the line interval [0,1], hence their sum cannot exceed the value of 1.

    Ordinary scaling of l_{ij} could be imposed such that its value cannot exceed the value of one.

13. In order to satisfy property 12, the coefficients in the matrices will have to be normalized in some defined manner.

14. a_{ij}^n coefficients may be fuzzy subsets of $(a_{ij})_{max}$, where $n \geq 0$. Thus, the a_{ij} coefficients may acquire n-tuple values before reaching their maximum value during the learning process. The n refers to the n th value that a_{ij} has acquired during the learning process.

    The augmented values during this process may occur in a manner described by equation 3.4.

15. $a_{ij}b_{ij} = \max \{a_{ij}^n, b_{ij}^m\}$. This defines specifically the fuzzy multiplication rule for the coefficients. The implication being that the resulting coefficient may be fuzzy subsets of the b_{ij} coefficients, and that several operations, m, of b_{ij} on a_{ij} may be required before it acquires the maximum value possible.

16. If $b_{ij} = 1$, then the product $a_{ij}b_{ij}$ defined in property 12 need not necessarily be satisfied due to the temporal aspect of the learning process, and property 15. Since several stages in the learning process

may be required, as was shown in figures 2.2 and 2.3. However a_{ij} may approach the value of one in some asymptotic manner.

17. $\frac{\partial a_{ij}^{n}}{\partial b_{ij}^{t}} \rightarrow (b_{ij})_{max}$, if $b_{ij} > a_{ij}$ The coefficients are interchangeable if $a_{ij} > b_{ij}$. This defines the important characteristic that the evolution of the coefficients, a_{ij}, during the learning process, depend on the effects that the b_{ij} coefficients have on them, and that they must tend towards them. The goal is for A to acquire the knowledge or understanding that B has. The t superscript refers to the number of times b_{ij} must operate on the a_{ij} coefficient in order to achieve learning or understanding, as in property 15. The n superscript refers to the subset of the a_{ij} coefficient. i.e. its value reflecting the state of knowledge at the time.
18. The evolution of the a_{ij} coefficients in the temporal context must also tend towards the b_{ij} coefficients, if $b_{ij} > a_{ij}$, thus $\frac{\partial^{2} a_{ij}^{n}}{\partial t \partial b_{ij}^{t}} \rightarrow (b_{ij})_{max}$

    This assumes that the b_{ij} coefficients are those of the expert and that learner A learns exclusively from expert B.

    The coefficients are also interchangeable if $a_{ij} > b_{ij}$.
19. $\sum_{i}\sum_{j} \alpha_{ij} \rightarrow U$ This is a more generalized statement referring to the fact that matrix of the learner will tend towards a state of knowledge during the learning process and is in consistency with equation 3.3. The α_{ij} are the normalized coefficients of the learner's matrix.

20. One must admit the matrix multiplication symbolized as $M_A \otimes M_A$. Whereby one considers multiplication of the elements such as $a_{ij} \circ a_{ik}$ ($j \neq k$). This implies the existence of a relationship between the a_{ij} coefficients, thus yielding the relational matrix R with coefficients described as e_{ij}^{ik}, $k \neq j$.

The characteristics and properties of the relational matrix would be similar to those of the learning matrix.

21. $L_i = \sum_{j=p}^{j=1} l_{ij}$ This defines the i th learning outcome for some specified primary concept, con (T_i) or procedure, proc*(a_{ii}).

The l_{ij} coefficients are related to the a_{ij} and b_{ij} coefficients via the fuzzy multiplication rule given by property 12, in which, $l_{ij} = a_{ij}\ b_{ij}$.

22. $l_{ij} \leq 1$. This must hold in view of the fact that the coefficients have may have fuzzy characteristics, as in property 5.

The above condition creates a situation where the coefficients will have to be normalized in view of the fact that property 7 is such that it otherwise would be possible that one may have the result that $L_k \geq 1$, and the fact that the fuzzy character must be preserved.

23. $U = \sum_{i=p}^{i=1} L_i$ This defines the final coefficient for the learning outcome or understanding . It could also be expressed as a column matrix, written as.

$$U = \begin{pmatrix} L_{11} \\ L_{22} \\ L_{33} \\ . \\ . \\ L_{nn} \end{pmatrix}$$

24. If the coefficients a_{ij} are independent of one another, then $e_{ij}^{ik} = 0 \quad (j \neq k)$. They can be said to possess the orthogonality property, in the geometric context, in which case they can be represented by a D dimensional hypercube, where $D = \sum_i \sum_j a_{ij}$ as defined in property 11.

Chapter 4 : Operationalization of the Matrices.

Normalization Scheme.

Conditions 10, 12, 13 and 14 along with the summation rule described earlier necessitate the application of some normalization scheme to the coefficients that appear in matrix equation, 2.3. This in view of the fact that, in the operational rules that were defined for these matrices, the coefficients may acquire values which exceed the numerical value of one. This situation would clearly be inconsistent with assigned fuzzy values, which by definition cannot acquire values greater than one since they are defined on the line interval [0,1], property 5.

A proposed normalization scheme for these coefficients can be adopted similar to the one proposed by Klir and Folger (1988), given by,

$$\alpha_{ij} = \frac{a_{ij}}{\sum a_{ij}} \quad and \quad \beta_{ij} = \frac{b_{ij}}{\sum b_{ij}} \qquad \ldots\ldots\ldots\ldots 4.1$$

One can also impose a similar scheme for the L_k coefficients, where,

$$L_K = \frac{l_{ij}}{\sum l_{ij}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.2$$

Thus, equation 3.3 with the normalized coefficients will have to be written using the set of α_{ij} and β_{ij} coefficients. The normalization scheme as expressed by equations 4.1 and 4.2 might be adequate under certain conditions, in particular those where the coefficients do not change significantly under the actions of some operations on them or with time. This however is not the case when applied to a learning process, since this is a dynamic process, by definition.

This can be easily illustrated in the case where learning is taking place. In this representation the a_{ij} and b_{ij} coefficients will acquire an increased numerical value, as both learn from each other, in the case of say a tutorial conversation. This will result in an increased value of their sums, i.e. the denominators of equations 4.1 and 4.2, which in turn will yield a decreased value of the normalized coefficients. The final parameter, L_K, will diminish in value which will indicate a decrease in learning. This would clearly be a contradiction to the manner in which this parameter is defined. Thus a new normalization scheme must be adopted, namely one which makes reference to the maximum value that the coefficients can acquire, that is,
where $(a_{ij})_{max} = (b_{ij})_{max} = 1.0$

In the case where person A learns exclusively from person B, the suggested normalization scheme would be one written as,

$$\alpha_{ij} = \frac{a_{ij}}{\sum_i \sum_j (b_{ij})_{\max}}$$

One needs to include the fact that these coefficients will acquire different values during the learning process, hence their values will only be valid at a given instant in time, property 14, thus one should include the n index in equation 4.1. The more complete and general equation will thus be,

$$\alpha^n = \frac{a_{ij}^n}{\sum_i \sum_j (b_{ij})_{\max}} \qquad \text{.................. 4.3}$$

Another normalization procedure that one can consider is the one in which reference is made to ones own state of knowledge. This can be considered in situations where A learns from B, or B from A, in which case there is no absolute frame of reference to work with. In addition one can monitor the progress of ones state of knowledge during the learning process. These normalized coefficients will be calculated according to the expressions,

$$\alpha_{ij} = \frac{a_{ij}}{\sum_i \sum_j a_{ij}} \quad \text{and} \quad \beta_{ij} = \frac{b_{ij}}{\sum_i \sum_j b_{ij}} \qquad \ldots\ldots\ldots\ldots\ 4.4$$

When an agreement of understanding occurs between them then the matrices should approach each others values, thus one can write, $U_A \Leftrightarrow U_B$ as in a similar manner to the expression $proc_A \Leftrightarrow proc_B$, in Pask's notation.

The U coefficients are calculated from the expressions,

$$U_A = \sum_i \sum_j \alpha_{ij} \quad \text{and} \quad U_B = \sum_i \sum_j \beta_{ij} \qquad \ldots\ldots\ldots\ldots\ 4.5$$

Equation 4.5 expresses a global coefficient of understanding as a totality of understanding of the variety of concepts that make up the topic that is learned. This is in effect the traditional grading system employed in most institutions. It is not inconceivable to partition the U coefficient within this matrix construct. The objective however will remain to obtain a global value.

Multiplication Procedures.

A specific matrix multiplication rule will have to be applied incorporating the possible fuzzy characteristics of the matrix elements. In general this will be written as the matrix equation,

$$M_A \otimes M_B = L \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ 4.6$$

Written fully in terms of their coefficients, the above equation is written as,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . & a_{1q} \\ a_{22} & a_{21} & a_{23} & . & . & a_{2q} \\ a_{33} & a_{31} & a_{32} & . & . & a_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{pp} & a_{p1} & a_{p2} & . & . & a_{pq} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} & . & . & b_{1q} \\ b_{22} & b_{21} & b_{23} & . & . & b_{2q} \\ b_{33} & b_{31} & b_{32} & . & . & b_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ b_{qq} & b_{q1} & b_{q2} & . & . & b_{pq} \end{pmatrix} = \begin{pmatrix} L_{1q} \\ L_{2q} \\ L_{3q} \\ . \\ . \\ L_{pq} \end{pmatrix} \quad \ldots\ldots\ldots 4.7$$

Where U, represents the learning outcome coefficient, yielding the column matrix, L_{ij}, as described in property 21. These are obtained from the specific matrix operational rules, defined as expressed by the following equation.

$$\begin{aligned}
&(a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + \ldots\ldots + a_{1q}b_{1q}) = (l_{11} + l_{12} + l_{13} + \ldots\ldots + l_{1q}) \\
&(a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23} + \ldots\ldots + a_{2q}b_{2q}) = (l_{21} + l_{22} + l_{23} + \ldots\ldots + l_{2q}) \\
&(a_{31}b_{31} + a_{32}b_{32} + a_{33}b_{33} + \ldots\ldots + a_{3q}b_{3q}) = (l_{31} + l_{32} + l_{33} + \ldots\ldots + l_{3q}) \\
&\quad . \qquad . \qquad . \qquad . \qquad . \qquad . \\
&\quad . \qquad . \qquad . \qquad . \qquad . \qquad . \qquad \ldots 4.8 \\
&\quad . \qquad . \qquad . \qquad . \qquad . \qquad . \\
&(a_{p1}b_{p1} + a_{p2}b_{p2} + a_{p3}b_{p3} + \ldots\ldots + a_{pq}b_{pq}) = (l_{p1} + l_{p2} + l_{p3} + \ldots\ldots + l_{pq})
\end{aligned}$$

The above can be written symbolically as,

$$\sum_{ij} a_{ij} b_{ij} = \sum_i \sum_j l_{ij} = L_k \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.9$$

<u>Comparison of the matrix construct with Neural Network.</u>

The representation described above has certain features that can be compared with representations of neural network. The coefficients of the matrix representing the learner, M_A Learning Matrix, LM, have a dual purpose of sorts. On the one hand, they represent the concepts or topics to be learned and on the second hand, they represent the state of knowledge acquired for the specific concept which they represent. The extent with which learning occurs will depends largely on the effects that the coefficients, b_{ij}, from the expert's matrix, M_B, have on the a_{ij} coefficients, along with the characteristics, λ_I, of the learner, characteristic 12.

It is possible to make a comparison, at the fundamental level, with a neural network, where one is concerned with input output signals. In this context the input signal is represented by the b_{ij} coefficients and the output signal is the learned coefficient l_{ij}, which is in effect the augmented a_{ij} coefficients, as described in figures 2.1 and 2.2.

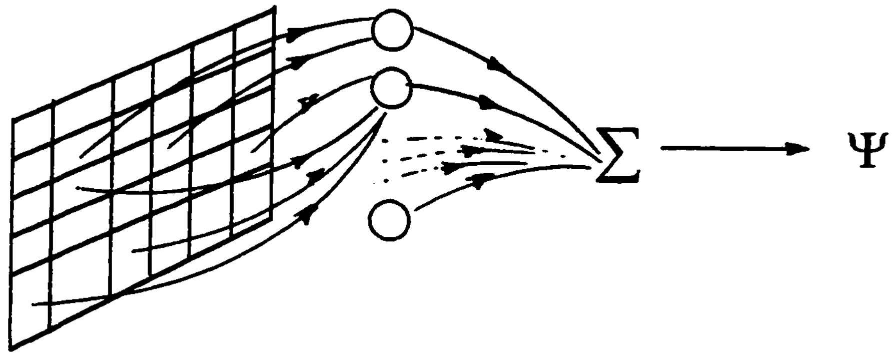
The difference between these two processes is that in the case of a neuron the relative importance of a synapse to the generation of a signal is the input coefficient, or the synaptic weight, ω_i. (Meaning that the neuron will not fire unless the input signal reaches a certain critical value). Whereas in the proposed matrix representation, the generation of a learning coefficient a_{ij} is not restricted to any minimum values of its input signals, the b_{ij} coefficients, and in principle, learning can always occur so long as $b_{ij} > 0$. i.e. the a_{ij} coefficients can always reach some value.

A further comparison can be made with Perceptron theory, again at the very fundamental level. The definition of a perceptron as given by Minsky (1990) is, " a perceptron is a device capable of computing all predicates which are linear in some given set of partial predicates."

A very simple description of a two layer perceptron given by (Kanerva, 1988), is shown in figure 4.1

Figure 4.1

Two layer Perceptron.



In this simple two layer perceptron, the first layer is a set of neurons with inputs from a "retina" that holds a (geometric pattern). The second layer is a (linear threshold) neuron that receives the input from the first layer. Its output is the output of the perceptron.

One of the interesting aspects of the neuron model is its potential for learning in a finite number of trials. This result is known as the " perceptron convergence theorem." In this respect it is analogous to the characteristics of the coefficients as given by 17, 18 and 14.

A further comparison, from the operational level, can be made with fuzzy systems, in particular FAM, (Fuzzy Associative Memory) matrices, Kosko (1988). In such systems, each entry defines an association between two parameters, in the FAM matrix, which defines a "rule" or an "input-output transformation". This architecture is described diagramatically in figure 4.2. However in the proposed matrix construct, the effect of the b_{ij} coefficients on the a_{ij} coefficients essentially act as the inputs and the resulting augmented coefficient values of a_{ij} are the outputs, which are in effect the l_{ij} coefficients given by characteristic 12, shown as a comparison in figure 4.3. These l_{ij} coefficients form fuzzy association quantities if the a_{ij} and b_{ij} are themselves fuzzy quantities, and equation 4.8 becomes a FAM (Fuzzy Approximation Matrix) type matrix. The FAM architecture as described by (Kosko, 1990) is shown in figure 4.2. The architecture of the learning matrix construct can be described in a very similar manner as shown in figure 4.3.

Figure 4.2

FAM Learning Architecture.

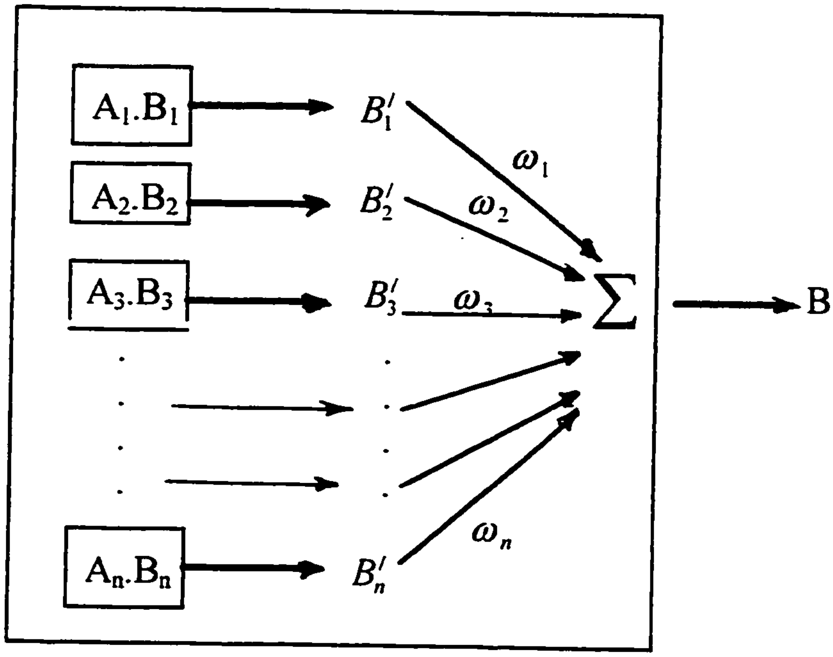
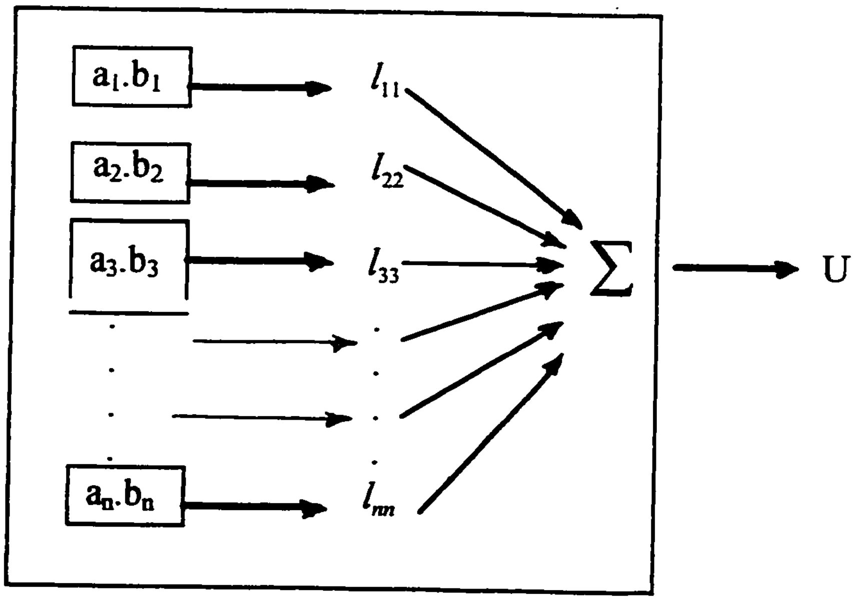


Figure 4.3

LM Architecture.



Evolution of the Matrices.

In the process of learning the objective is to reach or acquire a state of understanding. In the construct described previously, as a learning matrix, learning matrix, the concept map which this matrix attempts to describe, involves fuzzy elements since in general one can make the assumption that there is some understanding in existence, however its elements may be quite fuzzy in nature. The fuzziness of this matrix may be even more predominant if the topic under investigation is itself fuzzy.

As was stated earlier, the objective is to render this learning matrix less fuzzy. This means in effect reducing the fuzzy value of the matrix elements a_{ij}. Since the learning process takes place over some time interval then on can consider the evolution of the elements a_{ij} with respect to time thus introducing time differential terms of the form $\frac{\partial a_{ij}}{\partial t}$. This may only be necessary if one wishes to know this temporal evolution. Furthermore, since interaction between A and B involve the effects that coefficients of B have on the coefficients of A, then one can also consider the differential terms such as, $\frac{\partial a_{ij}}{\partial b_{ij}}$. This in general will not be constant and in certain restricted and well defined learning situations, it may be possible to determine that functional relationship.

The above refers to the fact that certain components, coefficients, in the learning process may be more difficult than others to reach the required level of understanding, hence their evolution will differ significantly as compared to those of lesser difficulty.

Much of the evolution of the a_{ij} coefficients with respect to the b_{ij} coefficients will be very dependent on the manner in which the conversation takes place, the modality of instruction or the method of teaching. By modality of the conversation it is meant the manner in which instruction takes place, in the pedagogical sense. It is possible to express the differential term described above in a manner which incorporates this modality.

One can thus write,

$$\frac{\partial a_{ij}}{\partial b_{ij}} \propto f(I) \qquad \text{4.10}$$

The above differential expression essentially states that the changes in the learning coefficients with respect to the known coefficients is dependent on the form in which instruction is given or the manner in which knowledge imparted to the learner. The form in which instruction is given, $f(I)$, may be the instructional design, that one chooses, or in general any pedagogical teaching modality chosen. The choice of $f(I)$ is highly context dependent. However in the absence of knowing which instructional design is the most suitable, it is possible to conduct experiments to determine the most suitable one. This would entail using some optimization technique.

The dependency on the characteristics of the student can be expressed generally and symbolically in an analogous manner as for the modality of instruction expression, that is,

$$\frac{\partial a_{ij}}{\partial b_{ij}} \propto g(\lambda) \qquad \text{4.11}$$

In addition, the characteristics of the learner, person A, will have to be taken into account, to the extent that this can be done.

One can thus write the evolution of the resulting learning matrix, as will be defined later, in terms of their coefficients, as a first order differential expression, and written as,

$$\frac{\partial M_A}{\partial b_{ij}} = \sum_i \sum_j \frac{\partial a_{ij}}{\partial b_{ij}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 4.12$$

The differential term, described above, is consistent with Pask's notion of Conversation in that it involves the evolution of the learner-teacher construct while the learning process is taking place. Thus the a_{ij} coefficients will undergo changes with respect to the b_{ij}, as expressed by property 17.

One can further generalize the matrix equation to include the modality of instruction, labeled I_i and the characteristic of the learner, labeled λ_i, by writing equation 4.6 as,

$$(M_B)^{I_i} \otimes (M_A)^{\lambda_i} = L \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 4.13$$

Thus I_1, I_2, I_3, ... I_i refer to different types of instruction, or modality of teaching. For example.

I_1 = conventional classroom setting, with lecture delivery from a teacher.

I_2 = teaching method relying entirely on overhead projection for lecture material dissemination and display.

I_3 = computerized learning method.

The λ_1, λ_2, λ_3, ... λ_i refer to the different types of learners. This would depend on the classification of students chosen. For example one can have,

λ_1 = deep learner. One wishing to grasp all the concepts to their full extent.

λ_2 = surface learner. One requiring an overall sound understanding of the topic to be learned.

λ_3 = superficial learner. One requiring but the minimum of understanding in order to pass a course to satisfy institutional requirements.

One can also express the time evolution of the matrix, globally as, $\frac{\partial M_A}{\partial t} \rightarrow M_B$. This again was defined specifically in property 17.

The term "understanding", U, can thus be defined as the sum of all the learned activities, L_i, or equivalently as the existence of all the necessary coefficients in the cognitive map, for person A, which can be written simply as, $U = \sum L_{ii}$ .... 4.14. This expresses property 21.

Given the evolution of the learning matrix as described above, then "understanding" must be treated as a dynamical state of learning, and is thus subject to continuous changes. The initial conditions of the a_{ij} and b_{ij} coefficients will be given by their fuzzy membership values attributed to them. The assignment of these values can be obtained from some initial pretesting condition on the subjects involved, or an assignment made by the subjects themselves.

It is not always possible to determine the functional relationship between the coefficients a_i and b_j. However, if a given instructional procedure has been employed on numerous occasions and its effect quantified in some manner, then it is possible to obtain some relationship between these coefficients.

The temporal relationship is more readily obtainable in view of the fact that one can administer a series of post-tests at some regular time intervals and measure the learning outcomes on some devised scale.

General Procedure.

The mathematical representation described in this paper for a learning process, can be more readily applied in the case of a one way conversation situation, i.e. a tutorial one. In this case the underlying assumption is that person B possesses all the knowledge that person A is to acquire. In this restricted case the above scheme is well suited for a learning situation whereby well defined knowledge is to be imparted on a specific task skill to be learned. The task may be a physical operational one such as : "how to operate a machine" or cognitive ones such as how to "solve a specific algebraic equation", "solve specific problems in Physics or Chemistry".

If the skill to be learned is a well defined one, then the components that make up the totality of the procedure, $proc_B^*(T_i)$, i.e the β_{ij} coefficients are regarded as mathematically crisp quantities. In the case where A learns or acquires all its understanding from B, i.e. a one way conversation, then the set of the β_{ij} coefficients will all have the value of one. The a_{ij} coefficients on the other hand will necessarily have fuzzy values, since $a_{ij} < 1$. For knowledge involving crisp entities one may question whether the a_{ij} coefficients can be fuzzy. This in the sense that either one knows how to perform a given task or one does not. However it is possible to encounter many situations in which one has partial knowledge of how to perform a certain task or at least some partial notions of what must be done. Thus, as an operational procedure of using these learning matrices one must follow the following steps :

1. Identify all the b_{ij} coefficients in an unambiguous manner.
2. Classify these coefficients in terms of their primary and secondary dimension in the matrix representing B.
3. Determine the order in which operations are to be performed .
4. Determine the initial values of the a_{ij} coefficients. i.e. their values at $t = 0$
5. Normalize the coefficients.
6. Select appropriate mode of instruction, for particular student characteristic
7. Perform the assigned operations on the coefficients.
8. Calculate the final learning outcome coefficient, U.

In order to operationalize these matrices, the fourth step is a critical one for much of what will develop in this system structure, will affect the final outcome. Recall that the goal of the differential expression, $\frac{\partial a_{ij}}{\partial b_{ij}}$ is to reach the values of the b_{ij} coefficients, for the case of person a learning from B. The time evolution of the learning matrix will also depend on certain characteristics of the learner, which can be labeled as $g(\lambda)$, where $g(\lambda)$ refers to some function relating to the characteristics of the learner, and also to some extent on the form of instruction, $f(I)$, or conversation taking place.

The complete learning matrix equation can thus be written in its most general form as,

$$\frac{\partial a_{ij}}{\partial b_{ij}} = \Omega f(I) g(\lambda) \quad \text{.......................... 4.15}$$

where Ω represents some constant, which may also be some function of time and related to $f(I)$ and $g(\lambda)$.

It must be pointed out that the order in which the terms in the matrix equation are used, in the operational sense, need not necessarily begin with the first element and continue in the order in which they appear, i.e. begin with the lowest value of the coefficients and proceed in their ascending numerical order, but rather on the basis of some defined context dependent procedural rules. Hence that procedural order will largely depend on the subject matter and the judicious choice of person B. If person B happens to be a computer, as in the case of Computer Assisted Learning, CAL, the order will be well defined and linearized.

Chapter 5 : Validation.

Applicability of the Representation.

This mathematical representation of a learning situation using concept matrices, as proposed, could be readily applied in a variety of situations. Some of these applications are described below.

1. *Measure of the knowledge state of a given individual.*

The coefficients in the concept matrix can be utilized to represent the measured extent of learning on a given concept at a given time. By evaluating this matrix it is possible to define an "understanding" parameter with respect to all the concepts. This would yield a global measure of understanding of the body of knowledge or topic that is tested. It is understood that the matrix evaluated, M_A, is as a result of the operation with matrix M_B.

It is also possible with this scheme to determine the time dependent characteristic of the state of knowledge. This in turn would enable one to obtain temporal learning curves, either for an individual or a group of individuals.

2. *Measure of instructional models.*

This in essence would measure the effects that different instruction strategies or teaching methods have on individuals or a group of individuals in terms of the knowledge imparted as measured by appropriate means.

In this situation matrix M_A is the one held constant in the sense that matrix M_B is varied and its effects on matrix M_A is measured. In procedural task instructions, the coefficients assume mathematically crisp values as opposed to fuzzy ones. The concept matrix M_B can thus be described unambiguously in addition to the order of operations to be performed.

The range of application can be from one requiring the learning of the operation of a mechanical device to one requiring the operational steps required to solve specific problems in disciplines such as Physics or Mathematics. From this it would be possible to determine the best or optimum teaching strategy or method which would yield the best or most efficient learning outcome.

3. *Measure of student models.*

The model of the student's knowledge can be constructed via the use of the concept matrices. In this context the coefficients of the matrix would represent certain characteristics of the student. One could construct a set of these matrices, each one representing a given characteristic. The elements of each of these matrices would then represent attributes or specific properties associated with that characteristic. One would then evaluate these matrices collectively, in some algebraic manner to obtain the "model" of that individual's tested knowledge.

It would then be possible to incorporate this model along with the concept matrix and determine learning outcomes. With this data it would be possible to adopt teaching strategies and methods which would maximize learning outcomes. That is, adapt teaching methods to fit the student model.

Validation.

The validity of the proposed scheme can and must be determined via two distinct perspectives. Firstly, the construct validity of the proposed representation must be determined since all subsequent applications rest on its defined procedures. Secondly, the results obtained from its application must also be validated.

a) Validation of the representation construct.

The first aspect to consider is the mathematical tool adopted for this representation. The criteria to be satisfied in this regard are that it be sufficiently flexible to itemize and display required entities and that one can devise rules that will enable one to manipulate the matrix coefficients appropriately. Matrices are quite common mathematical constructs and fit well for representing concepts in a systematic way, in view of the fact that they are fundamentally an array of symbols, labeled as coefficients. These coefficients can represent any desired entity. This construct permits the order and arrangements of these coefficients in any desired form. In particular, in the educational context, subject matter is often defined in an ordered structural manner in terms of its content, which may be a set of concepts to be learned. Matrices thus satisfy this requirement and hence can be valid for this type of description.

Once the matrix is adopted as a tool or descriptive device, then the operational rules associated with its coefficients must be examined within the context with which it is to be applied. Matrices, from the mathematical aspect have the characteristic that one can construct specific rules of operations. In the context of this representation this can be achieved provided that properties and characteristics of these coefficients are defined to correspond to the properties of tested knowledge. This was described in chapter 3. If these defined properties and characteristics show no contradictions or inconsistencies, then this criterion would be satisfied.

b) <u>Validation of its applicability.</u>

This second aspect of the validation requires two criteria to be satisfied, these being, firstly, the context within which this representation is applied must be appropriate, secondly, the results which this application yields are valid and meaningful. One can consider each in turn.

1. The context in which this representation will be applied is an educational one. The coefficients have been described as concepts and subconcepts associated with a body of knowledge to be learned. This was also described with an example, page 66. This example can be extended to other cases in the educational domain.

2. The results obtained must be valid and meaningful. This can be tested by considering different measures of performance utilized in the educational field. Since a normalization scheme was devised in order to render the numerical values meaningful, i.e. they will not reach values out of proportion, and the adopted scale of measure was the line interval [0,1] which in consistency with the standard scale. The validity of the measure itself needs also to be considered. This is discussed in more details in a subsequent chapter.

Formative Evaluation Procedure.

In order to implement the validation on the aspects of the proposed formalism, as discussed in the previous section, one can utilize some established criteria such as defined by Flagg (1990), who defined four kinds of formative evaluation performed during curriculum materials development for electronic technologies. Since this thesis deals with a formalism representing test item questions, it falls within this category of evaluation. The four kinds are :

1) *Connaisseur-based,* 2) *Decision oriented,* 3) *Object based,* 4) *Public relations-inspired studies.*

With respect to the proposed formalism the first category, the connaisseur-based, would be the most appropriate to apply for validation since it prescribes for expert in a given field to describe critically, appraise and illuminate the particular merits or demerits of a given object. These expert would examine the developed formalism with respect to the theoretical aspect of it as well as its subsequent applicability from their individual perspectives.

The evaluation requirements would be met by determining the following :

1. Correctness of applying matrices as constructs for knowledge representation.
2. Correctness of applying matrices to describe learner/ teacher interaction.
3. Correctness of operational rules devised within this construct.
4. The self consistency of the formalism.
5. Applicability to a wide range of educational subject matter.
6. Adequacy of determining specific learning outcomes on specific concepts.
7. Adequacy of representing state of knowledge.

A numerical example.

Person A learns from person B.

Consider the situation where a Physics classroom is taught the concepts and operations of a body moving in two dimensions.

It will be assumed that the modality of teaching is the conventional teacher-blackboard method. Following the procedures of the proposed scheme, the first step is to define the primary concepts required in order to acquire full understanding of this 2-D motion, these are the set of a_{ii} coefficients, and then the set of subconcepts associated with the primary ones, the a_{ij}, coefficients.

In this case, person B, the teacher may define these as :

a_{11} = vectors

a_{12} = pictorial representation

a_{13} = displacement vectors

a_{14} = velocity vector

a_{15} = acceleration vector

a_{22} = rectilinear motion

a_{21} = uniform motion, $v = \frac{dx}{dt}$

a_{23} = uniform acceleration, $a = \frac{dv}{dt} = \frac{d^2x}{dt^2}$

a_{24} = graphical representation of the relationship between the defined kinematics parameters.

a_{25} = general equations of linear motion

$$a = \frac{v_f - v_i}{t}$$

$$s = v_i t + \frac{1}{2}at^2$$

$$v_f^2 - v_i^2 = 2as$$

a_{33} = vertical motion

a_{31} = acceleration due to gravity

a_{32} = upward projections

a_{34} = downward projections

a_{44} = 2-D motion

a_{41} = angles of projections

a_{42} = resolution of velocity vectors

a_{43} = choice of equations

a_{45} = common variables

a_{55} = circular motion

a_{51} = direction change of acceleration

a_{52} = vector diagram

a_{53} = centripetal acceleration, equation $a_c = \frac{v^2}{R}$

The modality of instruction λ_i, is selected. Let this be a set of six, one hour lectures spread over a time period of two weeks, given by a teacher in a classroom lecture format. The type of the student can be taken into account from which the teacher decides on the level and depth of the content to be taught. It is assumed that there exists some homogeneity of the student population. This may be an extreme assumption to make given that the size of a classroom has typically 40 students, at the collegiate level. One will generally choose a level of the order of I_1. This to satisfy the surface learner.

The student concept matrix will appear as shown below, and the teacher's matrix will have the same number of coefficients with the numerical value of one for each, since it is assumed that he possess all the required knowledge. Matrix M_B also defines a structure of the concept matrix in terms of the coefficients that are required to describe it, correspondingly, matrix M_A will also acquire the same matrix form as expressed below.

$$M_A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{22} & a_{21} & a_{23} & a_{24} & a_{25} \\ a_{33} & a_{31} & a_{32} & a_{34} & \varphi \\ a_{44} & a_{41} & a_{42} & a_{43} & a_{45} \\ a_{55} & a_{51} & a_{52} & a_{53} & \varphi \end{pmatrix} \qquad M_B^{I_1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & \varphi \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & \varphi \end{pmatrix}$$

The dimensions are $D_P = 5$ and $D_S = 18$, thus $D = 23$. There are two nonexistent coefficients.

The initial state of knowledge of the student can be determined by administering a pretest. This would be done most efficiently on the terminal of a computer, since through the matrix representation the result can be displayed for each of the coefficients. Utilizing the line interval as the numerical scale, the initial, normalized and final, understanding matrix would appear as shown in table 1.

Table 1.

Numerical values of Coefficients and U.

matrix M_A					normalized matrix					L matrix
matrix at t = 0										
0.00	0.00	0.25	0.15	0.10	0.00	0.00	0.01	0.00	0.004	0.02
0.80	0.55	0.40	0.20	0.00	0.03	0.02	0.01	0.00	0.000	0.08
0.90	0.60	0.40	0.00		0.03	0.02	0.01	0.00		0.08
0.13	0.50	0.20	0.15	0.10	0.00	0.02	0.00	0.00	0.004	0.04
0.90	0.22	0.35	0.00		0.03	0.01	0.01	0.00		0.06
matrix at t = t1									**U =**	**0.30**
0.85	0.90	0.72	0.75	0.64	0.03	0.03	0.03	0.03	0.028	0.16
0.99	0.80	0.76	0.82	0.69	0.04	0.03	0.03	0.03	0.030	0.17
0.98	0.85	0.90	0.87		0.04	0.03	0.03	0.03		0.15
0.78	0.95	0.82	0.80	0.77	0.03	0.04	0.03	0.03	0.033	0.17
0.95	0.62	0.71	0.50		0.04	0.02	0.03	0.02		0.12
									U =	**0.801**

As one can note from the results displayed in the above matrices that the cases of $i_i < a_{ij}$ and $a_{ii} = 0$, $a_{ij} \neq 0$ at $t = 0$ can arise. In the example given above, $a_{ii} = 0$ but $a_{13} = 0.25$, $a_{14} = 0.15$ and $a_{15} = 0.10$

This can be explained in the following manner. A student may have no knowledge at all regarding a certain concept, however he or she may know something with regards to some of the subconcepts, although in an unrelated manner. However it is assumed that as learning progresses, property 4 will be satisfied. Thus at $t = 0$ the global understanding coefficient $U = 0.30$, after instruction, at $t = t_1$, that coefficient has acquired a value of $U = 0.80$. The above seemingly very simple example is but to illustrate the procedure. A more elaborate example could be devised to show how matrix M_A would evolve in time as the tutorial conversation unfolds, between A and B.

The numerical values of the coefficients will be derived from a test item bank of questions. The questions formulated are constructed so as to extract the degree of understanding for a specific component, or subcomponent of a given concept or topic to be learned. The constructed questions can be of two types :

a) Those that require numerical solutions to some specific problem, related to the concept in question, a_{ii} or a_{ij} .

b) Those that require a verbal responses.

It is possible to deal with linguistic variables through the use of fuzzy logic. Equation 4.8 defines the multiplication rules for the matrices, and the understanding coefficients were defined as, $U_A = \sum_{ij} \alpha_{ij}$ and $U_B = \sum_{ij} \beta_{ij}$ where α_{ij} and β_{ij} are the normalized coefficients. Using the computer as an operational tool it is possible to carry out the multiplication procedures of these matrices. Figure 5.1, which follows, is a flow chart showing how this can be actualized for the specific case of student A acquiring knowledge from person B. The essence of the operation is that a comparisons of the two matrices are made at every stage of the learning process. Thus , one begins at some time $t = 0$ and obtain the initial state of knowledge of the student, yielding matrix $(M_A)_{t=0}$.

A differential matrix, $(\Delta M)_{t=0} = (M_A)_{t=0} - M_B$, is obtained by considering the differences of all corresponding coefficients, thus,

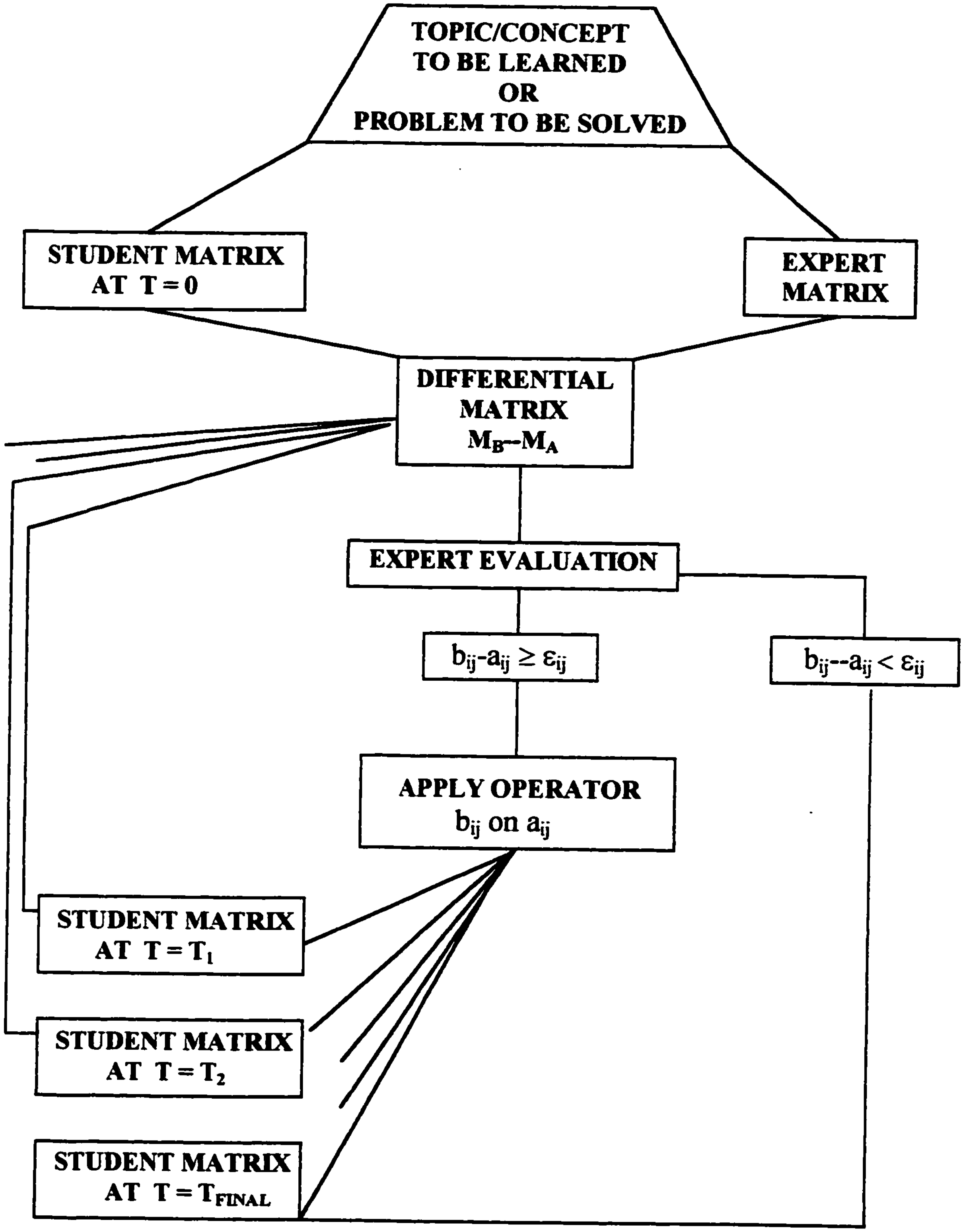
$$(\Delta M) = \begin{pmatrix} \beta_{11}-\alpha_{11} & \beta_{12}-\alpha_{12} & \beta_{13}-\alpha_{13} & . & . & \beta_{1q}-\alpha_{1q} \\ \beta_{22}-\alpha_{22} & \beta_{21}-\alpha_{21} & \beta_{23}-\alpha_{23} & . & . & \beta_{2q}-\alpha_{2q} \\ \beta_{33}-\alpha_{33} & \beta_{31}-\alpha_{31} & \beta_{32}-\alpha_{32} & . & . & \beta_{3q}-\alpha_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \beta_{pp}-\alpha_{pp} & \beta_{p1}-\alpha_{p1} & \beta_{p2}-\alpha_{p2} & . & . & \beta_{pq}-\alpha_{pq} \end{pmatrix} \quad \ldots\ldots\ldots\ldots 5.1$$

An evaluation is done by the expert with respect to every coefficient and a decision is made whether to proceed with the teaching or instruction for each component in the matrix. This decision can be based a chosen criterion. Taking into account the relative pre-requisite, importance and absolute value of each concept. In particular one can chose a boundary value for the differences, ε_{ij}. Hence if $\beta_{ij} - \alpha_{ij} \geq \varepsilon_{ij}$ this would signify that the knowledge difference between the expert and student is too large and hence learning must take place so as to augment the value of the coefficients. This process was in effect shown in figure 2.1

At every stage of the learning process the matrix M_A will acquire a new numerical value, since its coefficients have been augmented numerically. This would be described by the matrix M_A at different times $T = T_1$, $T = T_2$, $T = T_3$, ... , $T = T_{final}$. The set of l_{ij} coefficients that appear in equation 16 in effect represent those new α_{ij} coefficients. Correspondingly if this difference is less than ε_{ij} then the learning process is terminated.

Figure 5.1

Operational Flowchart for the Matrices.



A general case description.

In the following, a general case situation is described as this proposed formalism would, or, could be applied in a pedagogical setting. The case described here is one in which student A learns from an expert B. The case will first be described in a more generalized form and subsequently applied to a specific situation. This in view that the formalism of the proposed lerning matrix representation can be taken as a fundamental basis from which one can structure a particular application procedure. Although the application of this scheme is not meant to fall entirely within the ITS category, it will nevertheless employ many of its concepts.

The application proposed here is meant to be used as a secondary source of instruction and more importantly as a means of monitoring the state of knowledge with respect to the various concepts defined in the topic to be learned. The flow chart, as described in figure 5.2, (page 75) will form the basic architectural structure of this application and the following describe the procedural steps.

1. The set of primary concepts, a_{ii}, and corresponding subconcepts or subsets of them, a_{ij}, are clearly defined by person B since A's knowledge is taken to be a subset of B's.
2. Instructions/teachings are given to the student. This is represented as the operation of the b_{ij} coefficients on the a_{ij} coefficients. Thus, B explains or describes his understanding of the concepts.
3. With respect to each of the a_{ij} coefficients a set of questions are devised which will effectively test the knowledge acquired by the student. These set of questions can be labeled as Q_{ij}^{k}, where the k = 1, 2, 3, ... n

refers to the number of questions one wishes to assign in order to test the student's understanding of this particular a_{ij} coefficient. These will form a column matrix for each of the coefficients. The number of questions relating to each of the coefficients need not necessarily be the same. Thus n can vary for each coefficient a_{ij}.

4. The student now sits in front of the computer terminal and systematically answers all the questions posed. The outcome can be described in a corresponding column matrix. This column matrices will be subsets of of the a_{ij} coefficients. They can conveniently be labeled as the set a_{ij}^k.
5. An expert evaluation is now made with respect to each of the a_{ij}^k coefficients. This is done by making a comparison with the expert's answers. Thus the differences $b_{ij}^k - a_{ij}^k = \alpha_{ij}^k$ are calculated. Where the α_{ij}^k coefficients represent the knowledge state of the student at this particular stage of the learning process. These coefficients will have to be suitably normalized, as described earlier in the description of the formalism.
6. A criterion ε_{ij}^k is set for each of the coefficients.

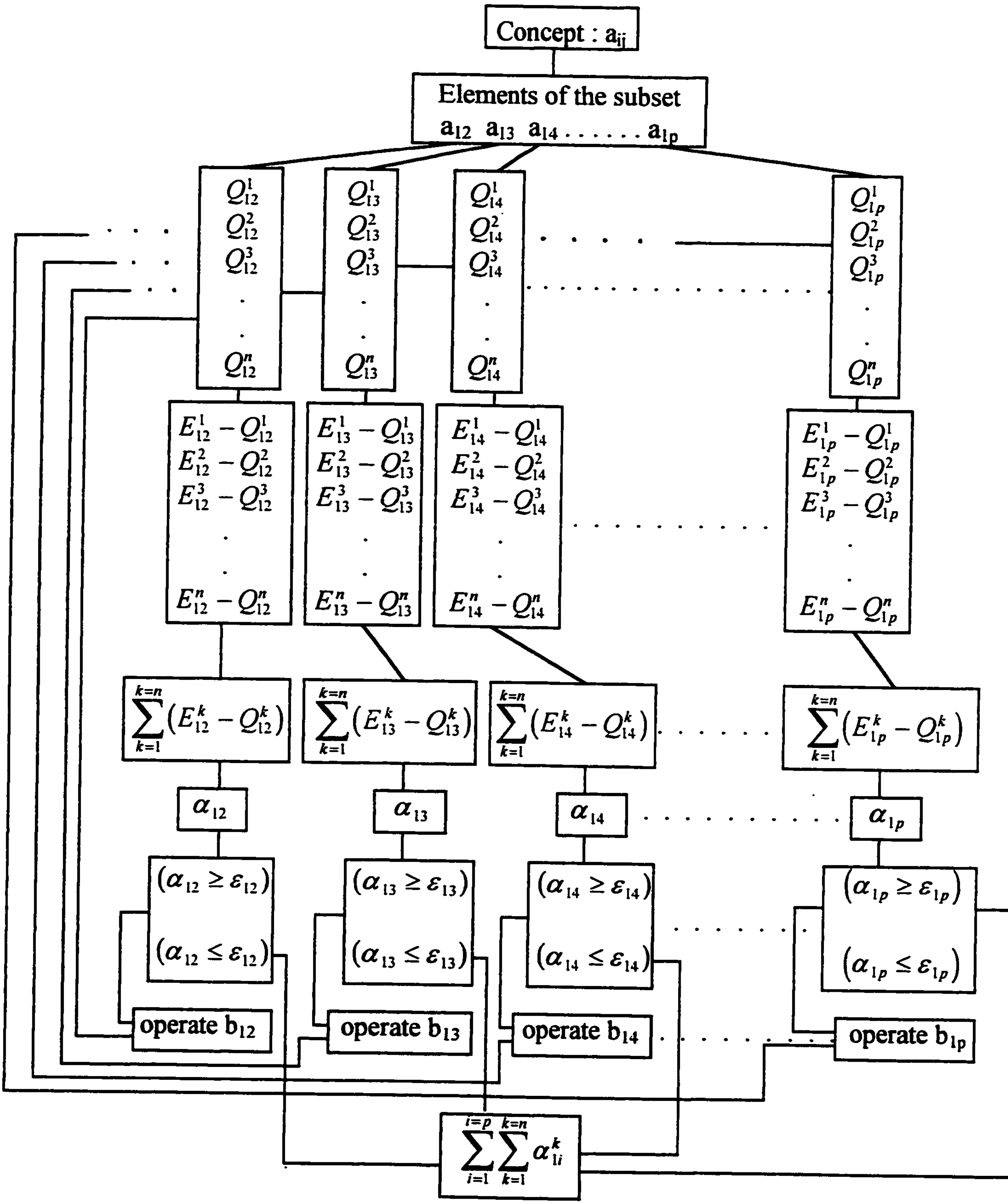
   If $\alpha_{ij}^k \leq \varepsilon_{ij}^k$ then the student is judged to have acquired sufficient knowledge in this particular element of the concept and his task is then terminated. If $\alpha_{ij}^k \rangle \varepsilon_{ij}^k$ this signifies that the gap between the expert's knowledge and that of the student is to large and the student requires additional instruction or practice to gain sufficient expertise in the concept. At this point the student is looped back for additional instructions and subsequently retested, however a different set of questions Q_{ij}^k will have to be used. Therefore, one will require several sets of the column matrices

Q_{ij}^{k} in this operational scheme.

7. At every stage of this process one can display the matrix with respect to the knowledge coefficients, and in addition one can calculate the global value of the coefficients, as given by the equation $U_A = \sum\sum \alpha_{ij}^{k}$

Figure 5.2 shows the "generalized operational flowchart" for the matrix.

Figure 5.2


Concept : a_{ij}
Elements of the subset
a_{12} a_{13} a_{14} ...... a_{1p}
Q_{12}^1 Q_{12}^2 Q_{12}^3 . . Q_{12}^n
Q_{13}^1 Q_{13}^2 Q_{13}^3 . . Q_{13}^n
Q_{14}^1 Q_{14}^2 Q_{14}^3 . . Q_{14}^n
Q_{1p}^1 Q_{1p}^2 Q_{1p}^3 . . Q_{1p}^n
$E_{12}^1 - Q_{12}^1$ $E_{12}^2 - Q_{12}^2$ $E_{12}^3 - Q_{12}^3$ . . $E_{12}^n - Q_{12}^n$
$E_{13}^1 - Q_{13}^1$ $E_{13}^2 - Q_{13}^2$ $E_{13}^3 - Q_{13}^3$ . . $E_{13}^n - Q_{13}^n$
$E_{14}^1 - Q_{14}^1$ $E_{14}^2 - Q_{14}^2$ $E_{14}^3 - Q_{14}^3$ . . $E_{14}^n - Q_{14}^n$
$E_{1p}^1 - Q_{1p}^1$ $E_{1p}^2 - Q_{1p}^2$ $E_{1p}^3 - Q_{1p}^3$ . . $E_{1p}^n - Q_{1p}^n$
$\sum_{k=1}^{k=n}(E_{12}^k - Q_{12}^k)$
$\sum_{k=1}^{k=n}(E_{13}^k - Q_{13}^k)$
$\sum_{k=1}^{k=n}(E_{14}^k - Q_{14}^k)$
$\sum_{k=1}^{k=n}(E_{1p}^k - Q_{1p}^k)$
α_{12}
α_{13}
α_{14}
α_{1p}
$(\alpha_{12} \geq \varepsilon_{12})$ $(\alpha_{12} \leq \varepsilon_{12})$
$(\alpha_{13} \geq \varepsilon_{13})$ $(\alpha_{13} \leq \varepsilon_{13})$
$(\alpha_{14} \geq \varepsilon_{14})$ $(\alpha_{14} \leq \varepsilon_{14})$
$(\alpha_{1p} \geq \varepsilon_{1p})$ $(\alpha_{1p} \leq \varepsilon_{1p})$
operate b_{12}
operate b_{13}
operate b_{14}
operate b_{1p}
$\sum_{i=1}^{i=p}\sum_{k=1}^{k=n}\alpha_{1i}^k$

Student and Teacher characteristics.

In the application of the proposed scheme as described in the previous section the type of student or learner was not taken into account. However in order for effective conversation to take place between a learner and a teacher, it is desirable that both have an adequate model of each other, (Mitchell, 1993), in order to maximize the understanding between them. Thus if a teacher has a good model of the student, he can then construct his teaching method and style to suite the student in order to maximize learning. Correspondingly if the student has an adequate model of the teacher then an adjustment in learning can be made to suite the teaching style being used. As mentioned earlier the task of modeling the student is far from being trivial and easy to perform. At best one can construct some models in terms of categories of students identified from the results.

The central point of this section is how the a_{ij} coefficients, which represent the student learning state, vary with the b_{ij} coefficients, those representing this expert or teacher. The variation between these coefficients was expressed as a proportionality with respect to the student characteristics $g(\lambda)$ described by expression 4.11, in chapter 4. The question of how to determine these characteristics which will in turn describe the type of the student still remains unanswered. It is however possible to give a symbolic description of the coefficients of the matrix representing the student. If one writes the proportionality expression of 4.11 as an equality by inserting an appropriate constant, then one has the equation,

$$\frac{\partial a_{ij}}{\partial b_{ij}} = k_{\lambda} g(\lambda) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.2$$

In principle the above is solvable since one can write the coefficients as,

$$a_{ij} = k_\lambda \int g(\lambda) db_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.3$$

The above equation has a rather significant interpretation associated with it, namely that the values of the coefficients of the student's matrix will depend on how the characteristics of the student are related to the coefficient of the expert's or teacher's matrix. The conclusion being that for learning to take place the teacher's characteristics, which are related to its b_{ij} coefficients, must be interwoven with those of the learner.

If one incorporates the expression 4.10, which relates to the modality of instruction, which in turn can be related to the characteristics or model of the teacher, then one can write a similar equation, as,

$$\frac{\partial a_{ij}}{\partial b_{ij}} = k_I f(I) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.4$$

Solving for the a_{ij} coefficients as in the previous case, then,

$$a_{ij} = k_I \int f(I) db_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 5.5$$

Equation 5.5 has a more obvious interpretation associated with it in that the model of the teacher expressed by the manner in which instruction is chosen is related to its b_{ij} coefficients, by definition. Combining equations 5.3 and 5.4 then the generalized equation for the a_{ij} coefficients can be written as,

$$a_{ij} = \Omega \int f(I) g(\lambda) db_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots 5.6$$

The above equation follows naturally from equation 4.13, namely,

$$(M_B)^{I_i} \otimes (M_A)^{\lambda_i} = L$$

and has the more generalized interpretation in that the models of both student and teacher must be taken into account in order to maximize the value of the

a_{ij} coefficients. Equation 5.6 assumes that the functions $f(I)$ and $g(\lambda)$ are both continuous. While this may be the true from purely theoretical terms, it may be quite difficult if not impossible to actually determine these functions. From the applied context it would seem that these functions can be utilized if treated as discreet quantities, in which case the integral sign would have to revert to a summation sign. Equation 31 would the be written as,

$$\left(a_{ij}\right)_{\max} = \Omega \sum_{m} \sum_{n} f^{m}(I) g^{n}(\lambda) b_{ij} \quad \ldots\ldots\ldots\ldots \quad 5.7$$

In equation 5.7, the indices m and n have been inserted to account for different forms of instructions possible and for different existing learner's characteristics. This is also consistent with the I_i and λ_I indices that appear in equation 4.13. In theory there would be mn combinations to yield $(a_{ij})_{max}$, however in practice one would accept only those combinations which are compatible with one another, as justified by B.

Equation 5.7 is more readily interpretable, in that in order to obtain the maximum possible value of the coefficients $(a_{ij})_{max}$, in the matrix, one might have to resort to a variety of teaching methods or styles, m of them, and appeal to various characteristics of the learner, n of them. This in itself is a process that often takes place in a classroom setting, where a teacher will resort to different methods of explaining a given concept, such as a diagrammatic one, a symbolic one and then concrete exemplary one. Similarly the teacher would try to seek the different characteristics of the student by, for example, speaking softly, or gesturing more or apply a more theatrical form in the delivery of the lecture. With respect to equation 4.13 the matrices can be related to the factors of equation 5.7 by,

$$(M_B)^{l_r} = \sum_m f^m(l) b_{ij} \quad \text{and} \quad (M_A)^{\lambda_r} = \sum_n g^n(\lambda) b_{ij} \quad \ldots\ldots\ldots\ldots\ldots 5.8$$

By incorporating the characteristics of the learner and teacher, the concept learning matrices defined previously can thus be developed in a more detailed manner. With respect to the above, one can ask the following question :

Given a collection of students which must be taught the same subject matter within the same time interval, can one devise a set of teaching instructions to suit each of the categories of students so as to achieve equivalent learning outcomes ?.

Fundamentally one would desire any construct to adapt itself to a variety of student models, or at least designed to achieve this goal. Research in this area is substantive and elaborate computer programs have been written to render this possible. An example of one is SIFT (Self-Improving Fraction Tutor), Gutstein (1992). This is a production system with a tutor and a learning module which learns from its interactions with students who use it. The above discussion simply illustrates that it could be possible, in principle, to incorporate these factors into account within the proposed construct, although no attempt is made to describe how this would be done other than in some very general context.

General formulation.

Symbolically, if the element of understanding a_{ij} of some topic T_i as understood by a learner type l_r is given by $[a_{ij}]^{\lambda_s}$ where $S = 1,2,3,\ldots.s$ and $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots \lambda_s$ represent the different types of students, then if the learning outcome is the same then the values of these coefficients should also be the same. This signifies that the learning matrices must have some final invariant character associated with them. This invariance can only come from an evaluation of these matrices by the same expert or related to some accepted standard of truth. It is possible to describe in a generalized manner, as in figure 5.4, the process by which it is possible to implement learning processes, incorporating forms of instructions $f^m(I) = I_i$, to different models of students, $g^n(\lambda) = \lambda_i$. In practice the implementation would be possible through the use of a computer where the available software is sufficiently flexible to accommodate the different models.

The process would be as follows. Consider any coefficient, a_{ij}, in the defined matrix. The value it acquires may depend on the model or type of learner, λ_S, hence one can write this coefficient as $[a_{ij}]^{\lambda_S}$,. This model of student can exercise a choice of form of instruction, I_i, for that particular concept and be tested or queried through a set of questions, q_{ij}^k, $j,k = 1,2,3,.......,n$, and k represents the number of questions devised for testing the knowledge for that specific concept. The responses will be compare to those of the expert in a similar manner as described in the flow chart of figure 5.2 and the final numerical or fuzzy value obtained to yield the coefficients in that matrix. The flow chart associated with this process is shown in figure 5.3. The central feature in this representation is the column matrix as expressed by equation 5.9.

$$\left[q_{ij}^{k}\right]^{I_i\lambda_n} = \begin{bmatrix} q_{i1}^{k} \\ q_{i2}^{k} \\ q_{i3}^{k} \\ . \\ . \\ . \\ q_{in}^{k} \end{bmatrix}^{I_i\lambda_1 \text{---} I_i\lambda_n} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 5.9$$

The above notation is equivalent to the one used in the chart of figure 5.1. As in the scheme represented in figure 5.3, one assumes that a set of k questions are constructed to test the knowledge of concept a_{ij}, and a set of such questions are constructed so as to be compatible with the mode of instruction given, in this context this is said to be equivalent to the model of the teacher or instructor. In this architecture, the set of questions associated with the ith form of instruction, I_i, must be made available to all types of students. Thus a learner has the option of choosing the form in which the set of questions are constructed. It might well be argued that all the learners should be tested with respect to different forms in which knowledge is represented so as to yield a better representative value for the coefficients.

Figure 5.3
Flow chart incorporating I_m and λ_s.

$a_{11}\ a_{12}\ a_{13}\ \ldots\ldots\ldots\ a_{1p}$

INSTRUCTIONAL MODEL I_m — $I_1, I_2, I_3, \ldots\ldots I_m$

STUDENT MODEL λ_s — $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots, \lambda_S$

$$\begin{bmatrix} q_{12}^1 \\ q_{12}^2 \\ q_{12}^3 \\ \cdot \\ \cdot \\ q_{12}^n \end{bmatrix}^{I_1\lambda_1 \text{---} I_1\lambda_s} \quad \begin{bmatrix} q_{13}^1 \\ q_{13}^2 \\ q_{13}^3 \\ \cdot \\ \cdot \\ q_{13}^n \end{bmatrix}^{I_1\lambda_1 \text{---} I_1\lambda_s} \quad \begin{bmatrix} q_{14}^1 \\ q_{14}^2 \\ q_{14}^3 \\ \cdot \\ \cdot \\ q_{14}^n \end{bmatrix}^{I_1\lambda_1 \text{---} I_1\lambda_s} \quad \cdots\cdots\cdots \begin{bmatrix} q_{1j}^1 \\ q_{1j}^2 \\ q_{1j}^3 \\ \cdot \\ \cdot \\ q_{1j}^n \end{bmatrix}^{I_1\lambda_1 \text{---} I_1\lambda_s}$$

$$\begin{bmatrix} q_{12}^1 \\ q_{12}^2 \\ q_{12}^3 \\ \cdot \\ \cdot \\ q_{12}^n \end{bmatrix}^{I_2\lambda_1 \text{---} I_1\lambda_s} \quad \begin{bmatrix} q_{13}^1 \\ q_{13}^2 \\ q_{13}^3 \\ \cdot \\ \cdot \\ q_{13}^n \end{bmatrix}^{I_2\lambda_1 \text{---} I_2\lambda_S} \quad \begin{bmatrix} q_{14}^1 \\ q_{14}^2 \\ q_{14}^3 \\ \cdot \\ \cdot \\ q_{14}^n \end{bmatrix}^{I_2\lambda_1 \text{---} I_2\lambda_S} \quad \cdots\cdots\cdots \begin{bmatrix} q_{1j}^1 \\ q_{1j}^2 \\ q_{1j}^3 \\ \cdot \\ \cdot \\ q_{1j}^n \end{bmatrix}^{I_2\lambda_1 \text{---} I_2\lambda_S}$$

$$\cdots \qquad \cdots \qquad \cdots \qquad \cdots$$

$$\begin{bmatrix} q_{12}^1 \\ q_{12}^2 \\ q_{12}^3 \\ \cdot \\ \cdot \\ q_{12}^n \end{bmatrix}^{I_m\lambda_1 \text{---} I_m\lambda_s} \quad \begin{bmatrix} q_{13}^1 \\ q_{13}^2 \\ q_{13}^3 \\ \cdot \\ \cdot \\ q_{13}^n \end{bmatrix}^{I_m\lambda_1 \text{---} I_m\lambda_S} \quad \begin{bmatrix} q_{14}^1 \\ q_{14}^2 \\ q_{14}^3 \\ \cdot \\ \cdot \\ q_{14}^n \end{bmatrix}^{I_m\lambda_1 \text{---} I_m\lambda_S} \quad \cdots\cdots\cdots \begin{bmatrix} q_{1j}^1 \\ q_{1j}^2 \\ q_{1j}^3 \\ \cdot \\ \cdot \\ q_{1j}^n \end{bmatrix}^{I_m\lambda_1 \text{---} I_m\lambda_S}$$

$$\frac{1}{m}\sum_{r=1}^{r=m}\sum_{k=1}^{k=n}\left[E_{ij}^k - q_{ij}^k\right]^{I_r\lambda_1} \quad \frac{1}{m}\sum_{r=1}^{r=m}\sum_{k=1}^{k=n}\left[E_{ij}^k - q_{ij}^k\right]^{I_r\lambda_2} \quad \cdots\cdots \quad \frac{1}{m}\sum_{r=1}^{r=m}\sum_{k=1}^{k=n}\left[E_{ij}^k - q_{ij}^k\right]^{I_r\lambda_s}$$

$$\sum_j \left[a_{1j}\right]^{\lambda_1} \qquad \sum_j \left[a_{1j}\right]^{\lambda_2} \quad \cdots\cdots\cdots\cdots \quad \sum_j \left[a_{1j}\right]^{\lambda_S}$$

The generalized operational scheme shown in figure 5.4 allows for all teacher and student types to interact. In order to determine the knowledge outcomes. one could consider a simple averaging from all the instructional models. Thus the term, $\frac{1}{m}\sum_{r}\sum_{j,k}\left[E_{ij}^{k} - q_{ij}^{k}\right]^{I_r\lambda_i}$

represents the average from a set of m instructional models for the i the student model that appears in that flow chart. This would yield the coefficients $\left[\delta_{ij}^{k}\right]^{\lambda_i}$. to be defined later. from which one can then obtain the set of coefficients. $\sum\left[a_{1j}\right]^{\lambda_i}$ for the i th student type. However, this may not be an easy process to implement given the complexity of the situation that it would pose. One of the more obvious problems would be to determine all the different possible types and thereby deal with all the possible combinations. Hence if there are m instructional models and s student types then one would have to deal with $\frac{N!}{(N-2)!\,2!}$ combinations of $I\lambda$., where $N = m + s$. A way to simplify this situation is to define some set theoretical criteria in which one would group a set of common characteristics and thus establish select types falling under those groups. Thus modeling techniques would first have to be applied to achieve this.

Degeneracies.

A problem that might arise is the case of degeneracy. That is to say, the situation where one of obtains the same learning outcomes, or knowledge $.L'$. by applying different models of instruction $f^{m}(I)$ to a given model of learners $g^n(\lambda)$, or applying the same instructional models to a set comprising of different models of students.

These cases would yield the following equalities,

$$\sum_j \sum_k \left[E_{ij}^k - q_{ij}^k\right]^{I\lambda_R} = \sum_j \sum_k \left[E_{ij}^k - q_{ij}^k\right]^{I\lambda_S} \quad \ldots\ldots\ldots\; 5.10$$

and,
$$\sum_j \sum_k \left[E_{ij}^k - q_{ij}^k\right]^{I_m\lambda} = \sum_j \sum_k \left[E_{ij}^k - q_{ij}^k\right]^{I_n\lambda} \quad \ldots\ldots\ldots\; 5.11$$

and after sufficient number of iterations of the process to minimize the term $\left[E_{ij}^k - q_{ij}^k\right]$, as in the process described in figure 5.3, one would obtain the final set of normalized coefficients, $\sum_i \sum_j \alpha_{ij}^{I_m\lambda_s}$

There are two cases that can be considered separately.

a) *Instructional model I kept constant and Student model λ varying.*

This is the standard conventional approach taken by most educational institutions, whereby the same teaching method is applied to a collection of students which can range from a half a dozen to several hundreds.

The chart as shown in figure 5.4 would yield the set of terms,

$$\sum_s \sum_{k,j} \left[E_{ij}^k - q_{ij}^k\right]^{I\lambda_s} \qquad s = 1,2,3,\ldots\ldots\ldots$$

from which the set of learning matrices, LM, can be obtained for the different models of students. Applying this to all the coefficients representing the topic to be learned, then the set of LMs will be written as,

$$\sum_S M_A^{I\lambda_s} = \sum_s \sum_{i,j} \left[\alpha_{ij}\right]^{I\lambda_s} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\; 5.12$$

From the set of learning matrices generated by the above equation it is then possible to obtain the corresponding Understanding coefficients, U, as expressed by equation 4.14, $U = \sum L_{ij}$

The degeneracy that might occur in this case is not a problematic one. In fact, it is quite expected since in a classroom of students of various types, many will achieve equivalent understanding of a topic. This is usually indicated by the equivalent score that students will achieve on the same test.

Generating this set of learning matrices for fixed I and varying λ will enable the determination of the function $g(\lambda)$, as given by the expression 4.11. This may be useful in determining which type of student learns better on a fixed form of instruction or teaching method.

b) *Instructional model I varying and Student model λ remaining constant.*

This is the case where a well defined set of types of students are exposed to different instructional forms for some topic. This may arise in situations where it is desired to reinforce the understanding. This often arises in physical science disciplines where one would employ a theoretical approach to teaching a given topic and in addition apply a practical approach to the same topic. Again the chart in figure 5.4 would yield the terms,

$\sum_m \sum_{k,j} \left[E_{ij}^k - q_{ij}^k\right]^{I_m \lambda}$ m = 1,2,3,........ from which the set of LMs can be obtained in the similar manner as in the case above, this will be written similarly as,

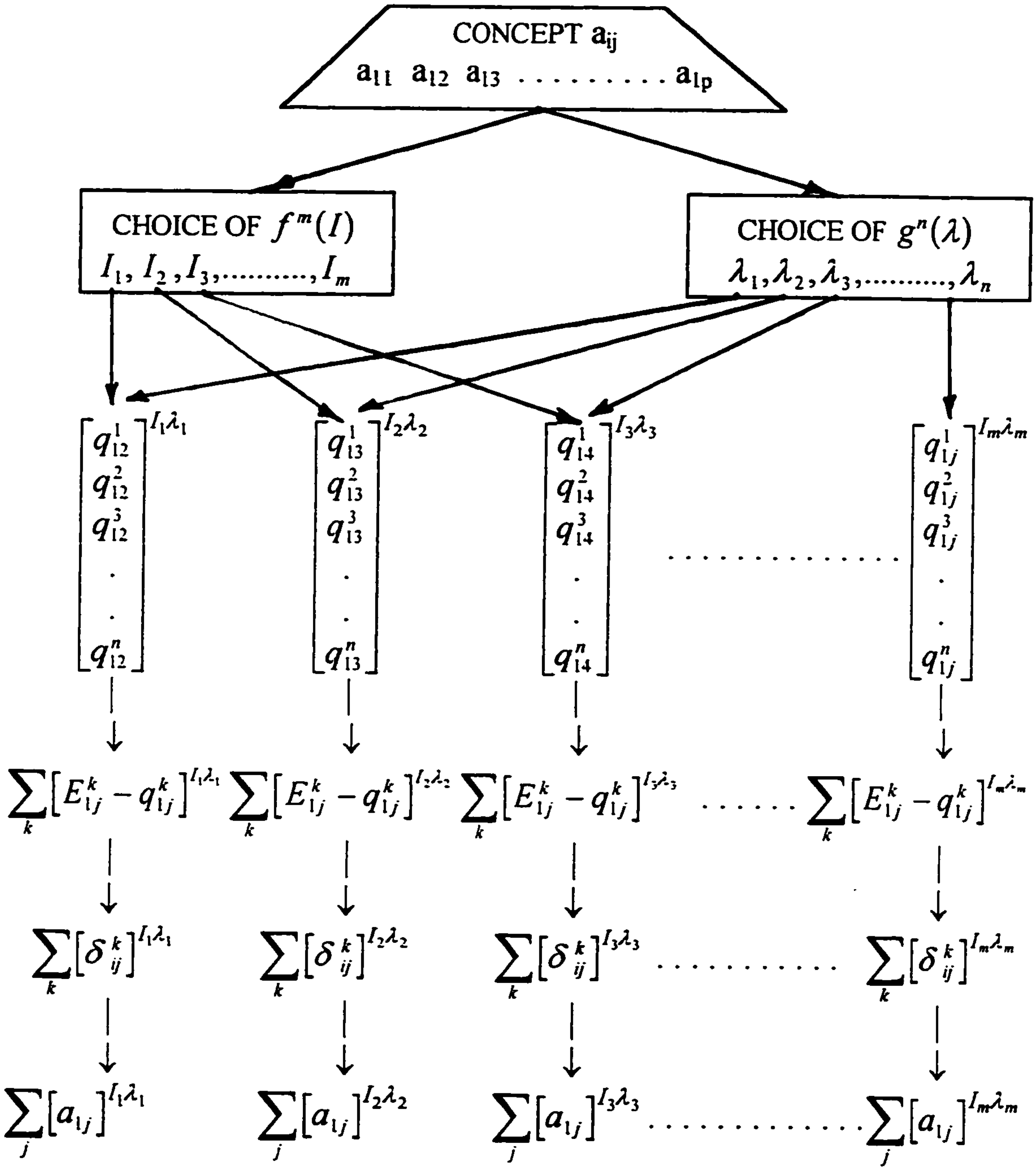
$$\sum_m M_A^{I_m \lambda} = \sum_m \sum_{i,j} \left[\alpha_{ij}\right]^{I_m \lambda} \qquad \ldots\ldots\ldots\ldots 5.13$$

The case of degeneracy as might arise from the set of learning matrices as generated by equation 5.13 may be more problematic in that one might question whether it is necessary to have different forms of instructions which yield the same learning outcomes.

One of the features of equation 5.13 is that it would be possible to determine the function $f(I)$ as given by expression 4.10. This would be extremely useful in determining which instructional form is most effective for a given type of learner. One way in which the redundancies of equations 5.10 and 5.11 can be resolved is by pairing the functions $f^m(I)$ and $g^n(\lambda)$ in such a way that they are pedagogically compatible. If one uses the same indices to indicate compatibility, then one can construct an operational scheme involving a new coefficient, (delta, δ). This coefficient will be described in detail in chapter 6 which follows. The flow chart involving these coefficients is shown in figure 5.4.

Figure 5.4

Flow chart incorporating the δ coefficient.



Chapter 6 : The Relational matrix.

Relationship between coefficients.

The coefficients that make up the learning matrices have the characteristics that they specify the pertinent subtopic to be learned within the global topic of interest or universe of discourse, and represent the knowledge state of the learner at a given time for that particular element. When structuring the learning matrix it is important to take into consideration that many of the coefficients may be related to one another. This could be from some knowledge domain subset characteristics that they may possess, or some common procedural learning characteristics. In general a relationship could arise depending on the specific properties of the coefficients. The strength of that relationship may not be easily quantifiable, if at all possible, nor may it be easily described. One should however make an attempt to determine the qualitative relationship that may exist between them. Failure to do so will diminish the richness of information associated with the learning matrices.

In work related to the level of knowledge acquired by learners, Frasson and Ramazani (1992), using the representation of inheritance systems, learning is described in terms of items of knowledge to be learned from an ensemble $K = [k_i , k_j , \ldots\ldots , k_n]$, with the idea that an element k_i inherits from an element k_j if one can find characteristics of k_i in k_j. In this sense a relationship between these items of knowledge is described. In addition, one defines the strength with which these inherited items are linked, termed "coherence". This enables the introduction of weight factors within this scheme. It then follows that in the matrix representation proposed here, one should give it a similar consideration.

The qualitative description can be rendered feasible by defining a set of qualitative properties and consider the relationship within the elements of that set. Jaworski and Grogono (1990) have devised a structural scheme which expresses how a set of elements are related to one another. They refer to their scheme as infoMAPS or J-maps. Initially one introduces the sets of interest, thus defining the universe of discourse. This is placed on a grid, such as a conventional spreadsheet, and entries are inserted which specify all the relationship between the elements of the defined sets, thus structuring the infoMAP. This scheme has the advantages that it can be used as a simple form of structural documentation of any topic of interest, which may quite useful in describing a course of study or as a structural description of specific learning processes required the understanding between elements of the sets. One can use the infoMAP matrix structure in conjunction with the matrix representation to specify and describe the relationship between the coefficients representing knowledge or learning procedures.

In order to obtain the relationship between the coefficients of a matrix one would require to perform a matrix multiplication of the matrix with itself, yielding the relational matrix, R. The multiplication rules however do not follow the standard matrix multiplication rules as described by property 20, page 47. This was expressed as,

$$R = M_A \otimes M_A \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 6.1$$

The resulting matrix referred to as the relational matrix R will have elements labeled ε_{ij}^{mn} where m, n, i and j represent the indices of the coefficients in matrix M_A. Each coefficient will have to be considered with respect to each and every other coefficient in the matrix. The following exemplifies the operations between the coefficients.

$$R = \begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . & a_{1q} \\ a_{22} & a_{21} & a_{23} & . & . & a_{2q} \\ a_{33} & a_{31} & a_{32} & . & . & a_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{pp} & a_{p1} & a_{p2} & . & . & a_{pq} \end{pmatrix} \otimes \begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . & a_{1q} \\ a_{22} & a_{21} & a_{23} & . & . & a_{2q} \\ a_{33} & a_{31} & a_{32} & . & . & a_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{pp} & a_{p1} & a_{p2} & . & . & a_{pq} \end{pmatrix} \quad \ldots\ldots\ldots 6.2$$

Considering the multiplication of elements in the i th row only for the above equation, this would yield a sub matrix r^i as,

$$r^i = \begin{pmatrix} a_{i2}^{i1} & a_{i3}^{i1} & a_{i4}^{i1} & a_{i5}^{i1} & . & . & . & a_{iq}^{i1} \\ . & a_{i3}^{i2} & a_{i4}^{i2} & a_{i5}^{i2} & . & . & . & a_{iq}^{i2} \\ . & . & a_{i4}^{i3} & a_{i5}^{i3} & . & . & . & a_{iq}^{i3} \\ . & . & . & a_{i5}^{i4} & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & a_{iq}^{i(q-1)} \end{pmatrix} \quad \ldots\ldots\ldots\ldots 6.3$$

If there are non zero coefficients in q columns and in p rows in the matrix, then the number of possible combinations of the coefficients will be $\frac{(pq)!}{[(pq)-2]!2!}$, and $\sum e_{ij}^{ij} = 0$, since the relation of a given coefficient cannot be considered with respect to its own self. In practice however many of these product coefficients will have zero values.

An example.

Consider as an example the area of knowledge to be learned as the Linear and Rotational Kinematics encountered in Physics. Procedurally, the Learning and Relational matrices are constructed in the following manner.

a) Identify all the elemental topics of interest and importance that constitutes major topic of "Kinematics" in the knowledge domain.
b) Group all the identified elements into their primary and secondary elements, in view of property 11, page 44.
c) Assign coefficient labels to these elements.
d) Construct the Learning Matrix, LM.
e) Determine what operations, relations and, or, procedures exist between these coefficients.
f) Construct the J-map as the architecture for displaying the relationship between the coefficients.
g) Construct the Relational Matrix, R.
h) Perform appropriate operations on the R matrix to determine the level or degree of understanding achieved by the learner. This as a result of operations of the corresponding R matrix of the expert or teacher.

The coefficients of the elemental topics of relevance for the example related to the concept of kinematics are described in table 2.

Table 2.

Description of the Coefficients.

a_{11} : **Rectilinear motion**

- a_{12} : vector representation
- a_{13} : linear displacement
- a_{14} : linear velocity
- a_{15} : linear acceleration
- a_{16} : S versus t graph
- a_{17} : S versus t^2 graph
- a_{18} : V versus t graph
- a_{19} : $a = \frac{v_f - v_i}{t}$
- a_{110} : $S = v_i t + \frac{1}{2}at^2$
- a_{111} : $v_f^2 - v_i^2 = 2aS$

a_{22} : **2-D Motion**

- a_{21} : acceleration due to gravity
- a_{23} : resolution of velocities
- a_{24} : $x = v_{ix} t$
- a_{25} : $y = v_{iy} t + \frac{1}{2}a_y t^2$

a_{33} : **Rotational Motion**

- a_{31} : angular displacement
- a_{32} : angular velocity
- a_{34} : angular acceleration
- a_{35} : angular acceleration
- a_{36} : $\alpha = \frac{\omega_f - \omega_i}{t}$
- a_{37} : $\theta = \omega_i t + \frac{1}{2}\alpha t^2$
- a_{38} : $\omega_f^2 - \omega_i^2 = 2\alpha\theta$

a_{44} : **Circular Motion**

- a_{41} : change in velocity
- a_{42} : vector diagram
- a_{43} : radial acceleration
- a_{45} : $a_c = \frac{v^2}{R}$

From the set of defined coefficients shown in table 2 one applies the J-map architecture, in a variant form, to describe and display the relationship between these coefficients. This is shown in table 3.

Table 3.
InfoMap.

InfoMAP FOR LINEAR AND ROTATIONAL KINEMATICS												
			LEGEND									
	F =	Figure		O =	Operation			d =	Definition			
	D =	Diagram		G =	Graphs							
	E =	Equation		M =	Many							
									MOTION			
				v				d	{RECTILINEAR MOTION}			
v		v								Vector representation		
F	v		v		v					Linear Displacement		
F	D	v				v				Linear velocity		
F	D	D	v				v			Linear acceleration		
	2			v				3		{GRAPHS}		
	G1										s versus t	
	G2										s versus t squared	
		G3									v versus t	
								3		{EQUATIONS}		
		E1	E1								s absent	
	E2		E2								final v absent	
	E3	E3	E3								t absent	
				v				d	{2-D MOTION}			
			D							Acceleration due to gravity		
										Angles of projections		
F		O	D							Resolutions of velocities		
	2							2		{EQUATIONS}		
	Ex										along X axis	
	Ey										along Y axis	
				v				d	{ROTATIONAL MOTION}			
	D				E4					Angular displacement		
		O				E5				Angular velocity		
			O				E6			Angular acceleration		
					2	2	3	3		{EQUATIONS}		
						E6	E6				angle absent	
					E7		E7				final w absent	
					E8	E8	E8				t absent	
				v				d	{CIRCUALR MOTION}			
		D								Velocity change		
F										Vector diagram		
			O							Acceleration		
			1					1		{EQUATION}		
			E9							centripetal acceleration		
				M					{EXAMPLES}			
				M					{PROBLEMS}			

The learning matrix, LM, from the description given in table 2 can thus be written as,

$$M_A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & a_{110} & a_{111} \\ a_{22} & a_{21} & a_{23} & a_{24} & a_{25} & a_{26} & 0 & 0 & 0 & 0 & 0 \\ a_{33} & a_{31} & a_{32} & a_{34} & a_{35} & a_{36} & a_{37} & 0 & 0 & 0 & 0 \\ a_{44} & a_{41} & a_{42} & a_{43} & a_{45} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \ldots\ldots\ldots 6.4$$

Using the infoMAP as constructed in table 3 one can now write the relational matrix as,

$$R = \begin{pmatrix} e_{13}^{12} & e_{14}^{12} & e_{15}^{12} & e_{23}^{12} & e_{42}^{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ e_{14}^{13} & e_{15}^{13} & e_{110}^{13} & e_{111}^{13} & e_{24}^{13} & e_{25}^{13} & e_{31}^{13} & e_{35}^{13} & e_{37}^{13} & 0 & 0 \\ e_{15}^{14} & e_{18}^{14} & e_{19}^{14} & e_{110}^{14} & e_{23}^{14} & e_{32}^{14} & e_{35}^{14} & e_{37}^{14} & e_{41}^{14} & 0 & 0 \\ e_{19}^{15} & e_{110}^{15} & e_{111}^{15} & e_{21}^{15} & e_{23}^{15} & e_{34}^{15} & e_{35}^{15} & e_{36}^{15} & e_{37}^{15} & e_{41}^{15} & e_{45}^{15} \end{pmatrix} \quad \ldots\ldots\ldots\ldots 6.5$$

The above describes, in a matrix representation, the relationship between the coefficients. If these coefficients represent processes to be learned or to be performed then it is possible to assign numerical values to them in order to quantify the state of understanding.

The important aspect of the relational matrix is that it can replace the original matrix and therefore would share the same properties and characteristics if it were to be applied in the context of a learning process. It is therefore taken that the expert, or person B in the tutorial conversation, would have a similar matrix representation with maximum numerical values for its coefficients. The relational matrix becomes a much richer entity since it refers relational knowledge as opposed to isolated concepts and one could argue that this aspect of knowledge is more relevant and important.

Characteristics of the relational matrix.

One can regard the learning matrix M_A, as expressed by equation 2.3, as a pq = n dimensional vector with binary components. Depending on the context for which this matrix is used, these vectors can represent points, patterns, addresses, words, memory items, data, events or procedures.
The R matrix then simply describes the relationship that may exists between these components and hence can serve the purpose of :

a) *Specifying the relationship as a purely declarative representation.*
In this context, the matrix merely serves as a descriptive device for representing a learning topic with respect to all the subtopics associated with it. The R matrix will then describe a verbal relationship, i.e. similarities of the descriptive words used that may exist between these subtopics. Although this may appear as trivial, it does have the advantage that given the systematic and well ordered structure it may facilitate understanding of the description given. More importantly it is a stage in the operational procedure of constructing the R matrix.

b) *Specifying the strength that exists between any relationships.*
It may be useful to know the strength or the degree to which the coefficients in the matrix are related to one another, the reason being that it may affect the modality of teaching or instructional methods chosen. Numerical values can be obtained, for these coefficients through similarity measurement techniques or fuzzy logic methods.

c) *Specifying operational procedures that define these relationships.*

In many cases, in particular in the context of education, the relationship between the coefficients may be such that some kind of operation or procedure may be required to be performed in order to reach an understanding of another concept.

In such cases the R matrix will be a much richer and more meaningful entity, and will in itself become a learning matrix since the operational procedures require understanding within their own sphere.

Characteristics.

1. A relational matrix R occurs by virtue of the existence of a relationship between at least one coefficient and any other coefficient in that matrix.
2. There can be only one defined relationship between any two coefficients. The coefficients e_{ij}^{ik} describe a one to one relationship between a coefficient a_{ij} and a_{ik} where $j \neq k$.
3. If each and every coefficient in an LM, with even number of coefficients, (a special case), has a relationship with only one other coefficient, the relational matrix will be a pure and a reduced one, since $\sum(a_{ij})_{coeff} > \sum(e_{ij}^{ik})_{coeff}$, expressed as, R_P^R .
4. If the LM has an odd number of coefficients the relational matrix will be a mixed matrix, written as R_M. It will therefore have elements e_{ij}^{ik} and a_{pq}.
5. If the coefficients in the LM have a many to one relationship with other coefficients, the resulting relational matrix will be an augmented one, expressed as R^A, if $\sum(e_{ij}^{ik})_{coeff} > \sum(a_{ij})_{coeff}$.

6. It may be possible, in principle, to reduce relational matrices to smaller reduced matrices. This is could occur if the topic to be learned can be represented in a well defined systemic structure, as for example in the physical sciences.

Operationalization of the InfoMaps.

The learning matrices are initially written in a manner which simply describes a topic to be learned in terms of its primary and secondary concepts and expressed by equation 6.4 in the example used. In applying the infoMap of Jaworski, in a variant form, it was shown that a relational matrix can be constructed, equation 6.5, which shows the existence of some relationships between the coefficients in the matrix. Since these coefficients can represent a process or procedure to be performed, it was subsequently argued that the state of understanding, U_R, should be of a higher order of magnitude as compared to that derived from simply U.

It is important to specify, in a manner as precise as possible, the exact nature of the coefficients in that constructed relational matrix, RM. To achieve this one can return to the initial infoMap and now write it in the format of an $n \times n$ spreadsheet array. The use of a spreadsheet is one of the architectures that can be used which happens to be a simple one to use for learners. Using the example of chapter 6, the relational matrix of equation 6.5 appears as shown in table 4. Note that the spreadsheet format is a matrix in itself. So that any knowledge represented in a spreadsheet can easily be subjected to the new matrix algebra diagnosis and prescription method of this thesis.

Table 4.

Spreadsheet Coefficient Display.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{110}	a_{111}	a_{21}	a_{23}	a_{24}	a_{25}
2		e_{13}^{12}	e_{14}^{12}	e_{15}^{12}										
3		e_{13}^{15}	e_{14}^{13}	e_{15}^{13}					e_{110}^{13}	e_{111}^{13}			e_{24}^{13}	e_{25}^{13}
4			e_{14}^{15}	e_{15}^{14}			e_{18}^{14}	e_{19}^{14}	e_{110}^{14}			e_{23}^{14}		
5								e_{19}^{15}	e_{110}^{15}	e_{111}^{15}	e_{21}^{15}	e_{23}^{15}		

	O	P	Q	R	S	T	U	V	X	Y	Y
1	a_{26}	a_{31}	a_{32}	a_{34}	a_{35}	a_{36}	a_{37}	a_{41}	a_{42}	a_{43}	a_{45}
2									e_{42}^{12}		
3		e_{31}^{13}	e_{32}^{13}		e_{35}^{13}		e_{37}^{13}	e_{41}^{13}			
4											
5											

As can be noted, the display of the coefficients as they would appear on the spreadsheet differs from that on the matrix representation, as given by equation 6.5. This difference is due to the operational function of the spreadsheet architecture. The spreadsheet architecture is executable in that each entry in each cell either describes a situation, a concept or a topic of knowledge, or a procedure to be performed. The entity in a given cell will in many of the cases make reference to entities in other cells by either assuming knowledge described in other cells or make direct reference to procedural steps described in other cell. The entries in the cells shown in table 4 have the following operational functions. (Note that the relational coefficients do not commute, thus $e_{ij}^{ik} \neq e_{ik}^{ij}$).

Cell A1: a_{12} - vector representation.

This would define concept of vectors. The choice of definition would depend on the pedagogical context and the level of complexity. As an example,

a) The existence of the set of solution (of a system of homogeneous linear equations, which a nonhomogeneous system can be reduced) consists of n-tuple of numbers : $\chi = (x^1, x^2, ...x^n)$, has a definite algebraic structure in which one can construct linear combinations. Such a structure is called a *vector space* and its elements are *vectors*, Behnke (1986).

b) A vector is a tensor of first rank, this has three components transforming as, $T_i' = \sum_j a_{ij} T_j$, Goldstein (1965).

c) A vector is a quantity which specifies magnitude and direction, Fowles (1977).

Cell B1, a_{13}, **Cell C1**, a_{14}, **Cell D1**, a_{15}, : describes concepts of linear displacement, velocity and acceleration respectively, thus,

a_{13} - Displacement. Net change of position of a body with respect to a Cartesian frame of reference.

a_{14} - Velocity. Change of position of a body with respect to a Cartesian frame of reference over a specified time interval. $v = \frac{\Delta S}{\Delta t}$, $\Delta S = S_f - S_i$, $t_i = 0$

a_{15} - Acceleration. Change of velocity of a moving body over a time interval, within the Cartesian frame of reference. $a = \frac{\Delta v}{\Delta t}$, $\Delta v = v_f - v_i$, $t_i = 0$

Cell B2, e_{13}^{12}, **Cell C2**, e_{14}^{12}, **Cell D2**, e_{15}^{12}, : describe the concepts of linear displacement, velocity and acceleration, respectively, in terms of the concept of vectors, thus,

e_{13}^{12} - Displacement. Resultant distance of a body relative to some point of origin, given by its magnitude and direction. The direction being specified by the angle of that displacement vector with respect to some specified linear axis on the Cartesian coordinate system. Thus requiring two numbers to fully describe it. $\vec{S} = (S, \theta)$

e_{14}^{12} - Velocity. The magnitude of the speed of a moving body and its direction of motion, as related to the Cartesian frame of reference. Thus requiring two numbers to fully describe it. $\vec{v} = (v, \theta)$

e_{15}^{12} - Acceleration. The magnitude of the change in velocity taking place over a time interval and the direction in which this change is taking place. Thus requiring two numbers to fully describe it.

Cell C3, e_{14}^{13}, **Cell D3**, e_{15}^{13}, **Cell C4**, e_{14}^{15}, **Cell D4**, e_{15}^{14}, : Describe the mathematical functional relationship of velocity and acceleration, thus,

e_{14}^{13} - Velocity, rate of change with respect to displacement, $v = \dfrac{ds}{dt}$.

This assumes displacement is expressed as a function of time, $S = f(t)$

e_{15}^{13} - Acceleration, rate of change with respect to displacement, $a = \dfrac{d^2 s}{dt^2}$.

This assumes displacement is expressed as a function of time, $S = f(t)$

e_{14}^{15} - Acceleration, rate of change with respect to velocity, $a = \dfrac{dv}{dt}$.

This assumes velocity is expressed as a function of time, $v = g(t)$

e_{15}^{14} - Velocity as an integral function of acceleration, $v = \int a\,dt$

This assumes acceleration is expressed as a function of time, $a = h(t)$

e_{13}^{15} - Displacement as an integral function of velocity, $s = \int v\,dt$

This assumes velocity is expressed as a function of time, $v = g(t)$

The operationalization of the infoMaps-maps thus become the operational process of the Rms. In the spreadsheet architecture, these matrices become "mixed" and "augmented" matrices, R_M^A. In the example described above the choice of information entered in each of the cells is subject to the decision of the expert or instructor. Quite clearly one can chose to amplify the descriptions of each of the entities entered by giving more details and associated examples, or enter numerical values to obtain solutions to particular kinematics problems.

Applying the spreadsheet as the operational mode renders this process more accessible to a wider range of learning audience at a variety of levels of learning depths. Furthermore, through the "window" computer environment it also becomes easier to use for the learners and it is possible to display the contents of the coefficients in the R_M^A for a variety of instructional modalities, I_i, discussed in chapter 5. This can simply be done by choosing a specific sheet within the spreadsheet application. It would also be possible, in principle, to cater to certain learner's characteristics through the use of different sheets for different characteristics. This is shown in figure 6.1.

Figure 6.1

Operationalization for chart of Fig 5.5.

Instruction I_1 student characteristic λ_1	Instruction I_2 student characteristic λ_2	Instruction I_3 student characteristic λ_3
sheet 1	sheet 2	sheet 3

The above figure thus shows how the scheme as described in figure 5.5 of chapter 5 can be implemented.

Associated learning coefficients.

The understanding, U, in the LM representation were obtained from the l_{ij} coefficients and were expressed by equation 4.8 as : $U = \sum_{i,j} l_{ij}$.

It is possible to obtain a similar expression for the understanding as would be derived from the relational matrix R. This can be achieved via the same procedure as adopted for the LM. Let U_R represent this coefficient, then,

$$U_R = \sum_{\alpha}\sum_{\beta} e_{\alpha}^{\beta} \circ f_{\alpha}^{\beta} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6.6$$

where $\alpha = i,j$, $\beta = k,l$, e_{α}^{β} are the relational coefficients, and f_{α}^{β} are the corresponding coefficients of person B or the expert. The multiplication rule as applied to the above equation is the same as for the ordinary coefficients that appear in the learner's matrix M_A and that of the expert M_B. Thus only products of corresponding coefficients can be considered.

Let the acquired learning coefficient be given by ξ, then one can write the expression : $e_{\alpha}^{\beta} \circ f_{\alpha}^{\beta} = \xi_{\alpha}^{\beta}$ this being consistent with property 8.

It should be noted that ξ_α^β is equivalent to the augmented e_α^β coefficient that would be created when learning is achieved. Of interest is how the l_α and ξ_α^β are related to one another. Equation 6.6 can be written as,

$$\begin{aligned} U_R &= \sum_\alpha \sum_\beta (a_\alpha \circ a_\beta) \circ (b_\alpha \circ b_\beta) \\ U_R &= \sum_\alpha \sum_\beta (a_\alpha \circ b_\alpha) \circ (a_\beta \circ b_\beta) \\ U_R &= \sum_\alpha \sum_\beta l_\alpha \circ l_\beta \end{aligned} \qquad \ldots\ldots\ldots\ldots\ldots\ldots 6.7$$

hence $\xi_\alpha^\beta = l_\alpha \circ l_\beta$ since $\sum_\alpha \sum_\beta l_\alpha \circ l_\beta \neq \sum_\alpha l_\alpha \circ \sum_\beta l_\beta$

It is clear from the above expression that $\xi_\alpha^\beta > l_\alpha$ hence $U_R > U$

This can be validated on the grounds that in order to reach the state of understanding given by ξ_α^β one requires to understand the relationship between two concepts, which may require understanding of a procedure or process in addition to understanding the concepts themselves.

Whilst understanding concepts in an isolated context may be important, one could make a strong argument for the case that understanding the relationship between concepts is far more important and relevant. An example of this can be taken from the realm of Physics, in dynamics where one requires an understanding of the concepts of force and acceleration as separate entities. In addition one would require to reach an understanding of the relationship between them which involves not only the mathematical expression and its associated manipulation rules, but also the concept of mass that arises out of this relationship. It might further be argued that, from a neural network representation of the process of understanding, the only process of relevance is that of the relationship among entities, or in the pedagogical context, the relationship among concepts.

Weight factors.

Since $U_R > U$, as was postulated earlier, then it might be of interest to determine the equality relationship between them. This can thus be written as, $U_R = wU \quad (w > 1)$ 6.8

The above equation leads to the notion of a weight factor, w. The question that needs to be addressed in this regard is where should these factors be introduced in the representation. Clearly these factors would be associated with the individual coefficients that appear in the matrix. The notion of weight factors is a natural one to consider and may indeed be a necessary inclusion in the matrix representation, since one would argue that when devising learning topics one invariably attaches more importance to some topics over others. It then follows that in measuring the state or level of knowledge one ought to take this fact into account. Therefore the weight factors need not be introduced in the initial learning matrices but rather in the subsequent matrices, L_k, of equation 4.9, these being the learning matrices that have evolved as a result of the operations of the b_{ij} coefficients on their respective a_{ij} coefficients.

The inclusion of the weight factors in this matrix representation is quite consistent with the FAM representation as described in figures 4.2, and also with Inheritance Systems described previously. Hence equation 4.9 would be rewritten as,

$$L_k = \begin{pmatrix} \omega_{11}l_{11} + & \omega_{12}l_{12} + & \omega_{13}l_{13} + & \cdots + & \omega_{1q}l_{1q} \\ \omega_{21}l_{21} + & \omega_{22}l_{22} + & \omega_{23}l_{23} + & \cdots + & \omega_{2q}l_{2q} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \omega_{p1}l_{p1} + & \omega_{p2}l_{p2} + & \omega_{p3}l_{p3} + & \cdots & \omega_{pq}l_{pq} \end{pmatrix} \quad \text{.........6.9}$$

$$U = \sum_{i,j} w_{ij} l_{ij} = \sum_{\alpha} w_\alpha l_\alpha \quad \text{......................6.10}$$

The value of w can acquire a meaningful value only with respect to the individual set of w_α coefficients.

It is possible to obtain the relationship between ww and w_α, w_β, by considering the case involving the relational matrix. From equations 6.7 and 6.8 one can write,

$$U_R = \sum_{\alpha}\sum_{\beta} w_\alpha w_\beta (l_\alpha \circ l_\beta) \quad \text{......................6.11}$$

therefore,

$$w = \frac{\sum_{\alpha}\sum_{\beta} w_\alpha w_\beta (l_\alpha \circ l_\beta)}{\sum_{\alpha} w_\alpha l_\alpha} \quad \text{......................6.12}$$

Equations 6.10 and 6.13 require that the weight factors be defined in some way as well as the product $l_\alpha \circ l_\beta$. In the simplest manner, the expert could assign numerical values in an arbitrary fashion to suite his or her sense of what the relative importance is of the coefficients in the LM. The numerical scale could also be one of personal choice. However, in the context of the teaching of a certain subject matter or topic, it is important that a universal standard consensus be arrived at with respect to the importance of the coefficients, since that is what these weight factors represent. This for instance would be of importance in an environment where multisectional courses are taught.

It is possible to obtain standardized values for these weight factors the details of which are described in appendix A. Assuming that these factors have been established, then, when applying equation 6.10 one must keep in mind that property 22 must be satisfied, that being that $U \leq 1$. If the numerical scale chosen for these factors is in the line interval [0,10], this would result in $U > 1$, in clear violation of property 22. This can be rectified by considering the weighted sums, then equation 6.10 would be written more appropriately as,

$$U = \frac{\sum_{\alpha} w_{\alpha} l_{\alpha}}{\sum_{\alpha} w_{\alpha}} \qquad \text{6.13}$$

similarly, equation 6.11 would become.

$$U = \frac{\sum_{\alpha} w_{\alpha} \sum_{\alpha} \sum_{\beta} w_{\alpha} w_{\beta} \left(l_{\alpha} \circ l_{\beta} \right)}{\sum_{\alpha} \sum_{\beta} w_{\alpha} w_{\beta} \sum_{\alpha} w_{\alpha} l_{\alpha}} \qquad \text{6.14}$$

With respect to the product of the learning coefficients that appear in equation 6.14, there are two possible operations that can be considered.

a) Applying fuzzy type multiplication rule, such as, $l_{\alpha} \circ l_{\beta} = \max(l_{\alpha}, l_{\beta})$.

This would be an easy operation to apply. However it would seem inappropriate to chose the maximum of the two since they may be mutually exclusive. Furthermore, understanding a set of concepts does not imply understanding another set.

b) Applying the standard Boolean multiplication rule. This would seem more reasonable since it may lead to a more "averaging" type value. In performing this type of operation one would clearly have to deal with the normalized values of these coefficients, that is, $\lambda_\alpha = \dfrac{l_\alpha}{\sum_\alpha l_\alpha}$.

The numerical values of the weight factors can be unambiguously determine through a mathematical procedure involving solutions of eigenvalue equations. The details are given in appendix A. One of the aspects of the procedure requires obtaining data from pairwise comparisons of the weight factors from a group of experts. The details of the procedure are given in appendix B.

Chapter 7 : Measurements of acquired knowledge.

The Delta coefficient.

In the procedural chart shown in figure 5.3 an important stage was the one involving the expert. This was expressed by showing the difference between the expert and the responses from the student/learner as $E_{ij}^{k} - q_{ij}^{k}$. This difference is important because it provides some numerical measure of the variance that might exist between the expert and the learner. This can be expressed by a single coefficient δ_{ij}^{k} defined by,

$$\delta_{ij}^{k} = E_{ij}^{k} - q_{ij}^{k} \qquad \text{7.1}$$

This coefficient will have properties depending on the nature of the set q_{ij}^{k}. This is due to the fact that q_{ij}^{k} can arise from questions involving strict computations or questions associated with linguistic variables. The distinction between the two being that in the former case one is dealing with logically crisp values which are associated with terms which are either right or wrong, true or false, in which case the determination of δ_{ij}^{k} is easily obtainable. The latter case involving questions which have vague or ambiguous notions will require linguistic terms such as quite weak, very strong, moderately good, etc. for their final response, and will thus require the use of fuzzy logic operational rules. One can thus make a distinction between the two case by writing this coefficient as, $\{q_{ij}^{k}\}_{N}$ and $\{q_{ij}^{k}\}_{F}$. Where N represents numerical questions, and F representing verbal question.

Relationships between $\{\delta_{ij}^{k}\}_{N}$ and a_{ij}.

Case i, for $\{\delta_{ij}^{k}\}_{N}$

From the definition as given by equation 5.3, the set $\{\delta_{ij}^{k}\}_{N}$ would acquire values ranging from 0 to 1. Thus, $0 \le \{\delta_{ij}^{k}\}_{N} \le 1.0$, where $\{\delta_{ij}^{k}\}_{N} = 0$ signifies complete agreement with the expert, or complete knowledge, and $\{\delta_{ij}^{k}\}_{N} = 1.0$ signifies complete disagreement or complete misunderstanding.

The next stage requires a summation of all these differences in order to make an assessment of the understanding or agreement reached, from which the α_{ij} coefficients can then be obtained. In the first instance it would appear that a simple arithmetic averaging would relate these two coefficients, such as $\alpha_{ij} = \frac{1}{n}\sum_{k}\delta_{ij}^{k}$ where n represents the number of questions to which the student must respond to. However, given the numerical properties mentioned above it is clear that this relation cannot hold for in fact the diametrically opposed values are related. Thus if $\{\delta_{ij}^{k}\}_{N} = 0$ then $\alpha_{ij} = 1.0$ and if $\{\delta_{ij}^{k}\}_{N} = 1.0$ then $\alpha_{ij} = 0$. In order to satisfy the above, the relationship can be written as,

$$a_{ij} = 1 - \frac{1}{n}\sum_{k=1}^{k=n}\{\delta_{ij}^{k}\} \qquad \dots\dots\dots\dots\dots\dots\dots \; 7.2$$

An example of the prototype of the set of numerical questions is given.

case i, $\{q_{ij}^{k}\}_{N}$

Consider the case of the concept of acceleration represented by the coefficient a_{22} as the primary concept within the global concept of kinematics that is to be taught. Suppose this has been taught in the traditional manner, say in the classroom, and it is desired to know the degree of understanding reached by the students. Following the procedure outlined previously, one needs to construct a subset of this primary concept. This subset then constitutes the conceptual representation of the concept acceleration. In this context we have the following,

a_{11} = acceleration.

a_{12} = Graphical representation.

a_{13} = Rates of change.

a_{14} = Linear equations of motion.

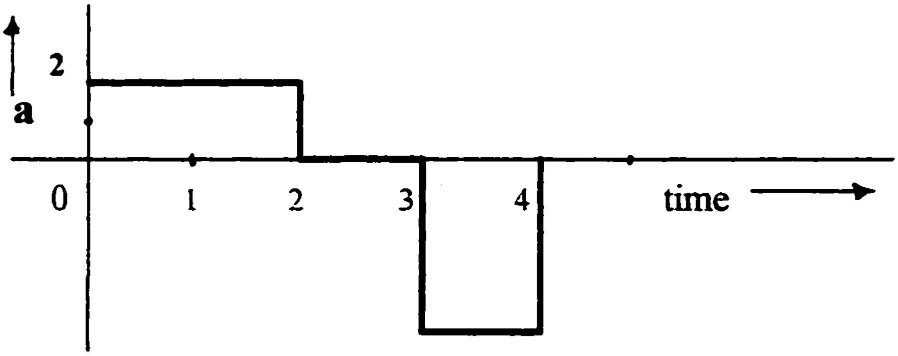
With respect to each element in the above subset, one devises a set of questions, Q_{ij}^{k} $k = 1,2,3,.....$, which will reveal the degree of understanding. If the questions are of the multiple choice type, which will be the case in this example, then their numerical values will be a crisp one as opposed to fuzzy values. The set of questions devised are :

Q_{12}^{1} = If the velocity versus time graph is a linear one, how would you determine the acceleration ?
a) read the intercept on the velocity axis.
b) calculate the area under the graph.
c) calculate the slope.
d) read the intercept on the time axis.

Q_{12}^{2} = If the velocity versus time graph is a horizontal line, how would you determine the acceleration ?
a) one cannot determine the acceleration.
b) calculate the area under the graph.
c) read the intercept on the velocity axis.

Q_{12}^3 = In an acceleration versus time graph, the area under the graph yields,

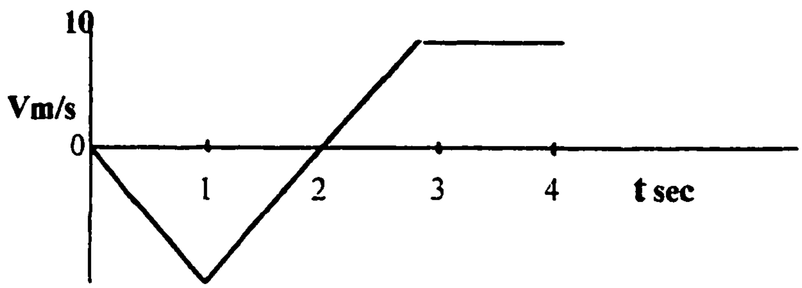
a) the velocity of the moving body.
b) the displacement of the body.
c) does not yield anything at all.



Q_{12}^4 = Calculate the velocity at the time of 10 seconds, and 20 seconds, if the initial velocity was zero.

Q_{12}^5 = Calculate the average velocity during the entire time interval of 25 sec.

Q_{12}^6 =



What is the average acceleration in the first 4 seconds, between the 2nd and 6th second, and in the 8th second

---

Q_{13}^1 = A particle moves along the X axis according to the equation $X = 2.0 + 5t^2$. Calculate its velocity and acceleration.

a) v = (2.0+10t) m/sec d) a = 5t m/sec^2
b) v = 5 m/sec e) a = 10 m/sec^2
c) v = 5t m/sec f) a = 10t m/sec^2

Q_{13}^2 = A particle has a velocity given by $v = 8t^2 + 3t + 5$, is this acceleration uniform ? find it's acceleration at a time t = 2 sec.

a) a = 35 m/sec^2
b) a = 19 m/sec^2
c) a = 40 m/sec^2

Q_{13}^{3} = A body moves according to $X = 5t^3 - 12t^2 - 6t - 4$. What are its velocity and acceleration in the third second ?

a) v = 57 m/sec d) a = 56 m/sec^2

b) v = 63 m/sec e) a = 66 m/sec^2

c) v = 135 m/sec f) a = 70 m/sec^2

Q_{14}^{1} = The initial speed of a body is 5.2 m/sec. What is its speed if it accelerates uniformly at 3.0 m/sec^2 and at -3 m/sec^2 ?

Q_{14}^{2} = A hockey puck sliding on ice comes to rest after traveling 200m. If its initial speed is 3.0 m/sec, calculate its acceleration, its speed after traveling 100 m and the time taken to travel the 200 m.

Q_{14}^{3}= A body travels along the positive X direction for 10 seconds at a constant speed of 50 m/sec, it the accelerates uniformly to a speed of 80 m/sec for the next 5 seconds. Find the average acceleration for the first 10 seconds and in the interval between the 10th and the 15th second.

Q_{14}^{4} = A body moving with uniform acceleration has a velocity of 12 m/sec when its coordinate is 3.0 m. If its coordinate 2.0 seconds later is -5 m, calculate its acceleration.

Case ii , $\{q_{ij}^{k}\}_F$.

While the proposed formalism is quite precise, that does not imply that its constituent coefficients are themselves precise and crisp quantities, for example the coefficients would have to be fuzzy if one were to apply this formalism to cases dealing with verbal descriptions and responses from learners. For such cases the evaluation of knowledge or assessment of agreement is somewhat more complicated since the responses would involves linguistic variables which have fuzzy properties and as a consequence the evaluation or assessment itself would also involve linguistic variables.

Typically the resulting assessment or evaluation, from modeling with some constructed production rules, might be that the student has :

a) not understood the topic very well.

b) has some understanding of the topic.

c) understood the topic quite well.

d) has a very good understanding of the topic.

The determination and quantification of such verbal ratings would necessitate the use of fuzzy membership curves. The determination of these curves may be initially tedious in certain respects, but once obtained they can be used continuously. Consider for example a case where a certain concept is to be described, the description of this topic may be such as to have to include a given set of words, as judged by an expert to be important and relevant to the full description of this concept. These may be the set $\{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}\}$. Fuzzy membership curves can be constructed to represent the "degree of understanding" from the fuzzy subset described above. These hypothetical membership value curves are shown in figure 7.1.

Alternatively, one can devise a set of fuzzy production rules that could be used for cognitive diagnosis, as in the case described for ACME (Abstract Modeling and Evaluation system.), Mitchell (1989). These could be of the form :

a) IF ALL of the above terms are used AND the order in which they appear offers NO contradiction AND no inconsistencies, THEN the understanding is VERY GOOD.

b) IF MOST of the terms in the set are used AND there are NO inconsistencies, THEN the understanding is GOOD.

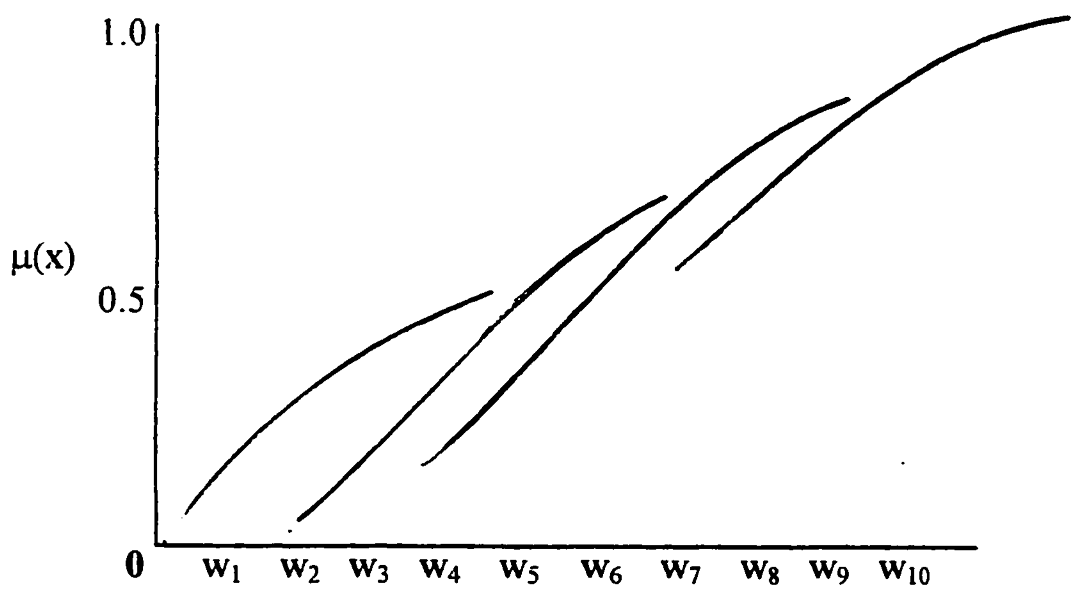
b) IF FEW of the terms are used AND no contradiction exists, THEN there is LITTLE understanding.

c) IF FEW of the terms are used AND there is inconsistency in their order of appearance, THEN there is VERY LITTLE understanding.

d) IF some of the terms are used AND there is contradiction in their appearance, THEN there is NO understanding.

Figure 7.1

Family of hypothetical membership curves.

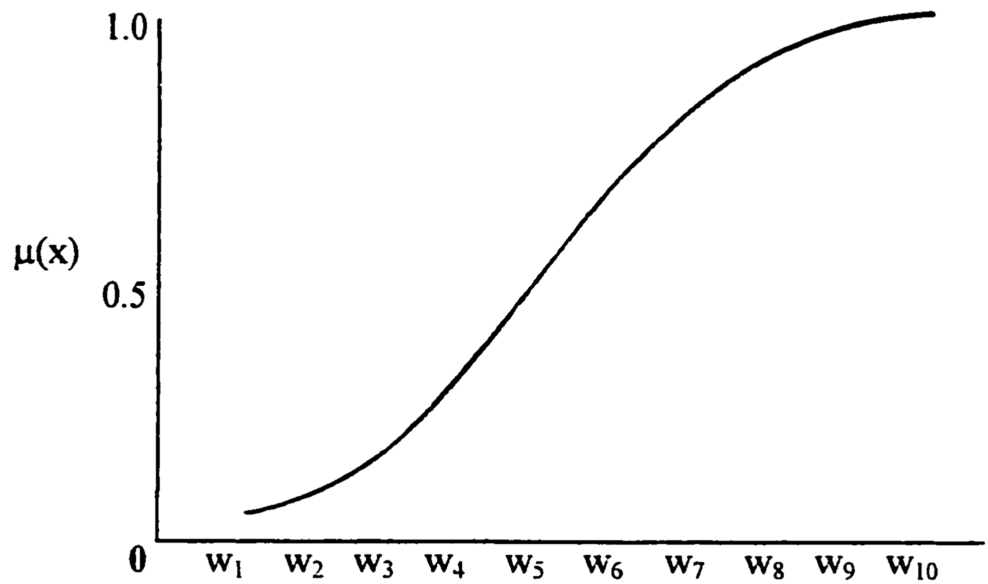


The set of curves shown in figure 7.1 would be adequate if detailed classification of verbal ratings were required, however, for the purposes of determining the a_{ij} coefficients of the matrix it is quite sufficient to simply establish one membership curve representing the term acquired knowledge and apply the numerical value of the membership $\mu(x)$ as the one which has direct correspondence with the a_{ij} coefficients. This however would imply that it would be necessary to establish a collection of membership value curves for each and every element in the matrix, the fuzzy set.

Suppose that one of these membership curves has been obtained and hypothetically is as shown in figure 7.2.

Figure 7.2

Hypothetical fuzzy membership curve.



The above curve resembles, geometrically, one of the collection shown in figure 7.1. This family of curves can be represented algebraically in a more general form than that as written by equation 1.2, namely,

$$U = \frac{1}{\sqrt{c_1 + \frac{D^2}{c_2}}} \quad \ldots\ldots\ldots\ldots\ 7.3$$

In order to establish the correspondence between the terms of the equation above and the ones defined previously it can be seen that the understanding symbol U represents the understanding for that particular concept, a_{ij} and hence could simply be replaced by that coefficient, and D effectively represents the degree of difference between the student's response and that of the expert's, that is, $D=\left[\delta_{ij}^{k}\right]_{avg}$, where $\left[\delta_{ij}^{k}\right]_{avg}=\frac{1}{n}\sum_{k=1}^{k=n}\delta_{ij}^{k}$.

The relationship between the a_{ij} coefficients, which represent the degree of understanding of the learner for the ij th concept and the δ_{ij}^{k} coefficient can now be written as

$$a_{ij}=\frac{1}{\sqrt{c_1+\frac{\left[\delta_{ij}^{k}\right]_{avg}^{2}}{c_2}}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ 7.4$$

Equation 7.4 can be used adequately to determine the coefficients of the matrix for situations involving linguistic variables, where in effect $\mu(x)$ corresponds to a_{ij} and the horizontal axis can be scaled in the interval $[0,1]$ to represent $\left[\delta_{ij}^{k}\right]_{avg}$. An example of descriptive type questions for the same concept, as in example i, is given by the below. Note that the correct answers to these questions must contain all the relevant parts, in terms of the correct words or phrasings, as for example the set of words $\{w_i\}$.

Example ii.

Q_{12}^{1} = How can average velocity be described ?, exemplify.

Q_{12}^{2} = Give a conceptual description of uniform velocity and exemplify it.

Q_{12}^{3} = Describe non uniform velocity and exemplify it.

Q_{12}^{4} = How would you describe the rate of change in kinematics, and give at least two examples of it.

Q_{13}^{1} = What are the fundamental quantities required in order to establish the equation defining acceleration ?.

Q_{13}^{2} = What are the distinctions between uniform and non uniform acceleration ?.

Q_{13}^{3} = Describe the motion of a light object falling freely in the earth's gravitational field.

Q_{14}^{3} = Can the equations of kinematics be used in situations where the acceleration varies with time or is equal to zero ? explain

Q_{14}^{3} = If there were three unknown kinematics parameters, can one use the three derived linear kinematics equations to solve for those three unknown ? explain.

Q_{14}^{3} = A person throws a ball upwards with an initial velocity v_o and another ball is dropped at the same instant, how would you compare the motion of these balls before they hit the ground ? explain.

Q_{14}^{3} = A ball is thrown upwards from the edge of a building of height h with an initial speed of v_o. A second ball is thrown downwards. How would their final velocities compare as they hit the ground ? explain.

Applied measures.

The coefficient representing the acquired knowledge for a given learning situation, for some learner A, was given by equation 4.5, expressed as, $U_A = \sum_{i,j} \alpha_{ij}$ where α_{ij} are the normalized values of the a_{ij} coefficients determined from either equations 4.3 or 4.4, depending on which case is applicable. The determination of a_{ij} in turn depend on the δ_{ij}^{k} coefficients, as defined previously, for which their relationship was given by equation 5.3, as exemplified in case i, if the question-response situation demanded crisp logical values, or equation 5.5 if it demanded fuzzy set logical values, as exemplified in case ii.

Since δ_{ij}^{k} is a critical value to be determined during the learning process, it is important to reconsider it in view of the manner in which the original notions of the matrix were used to define it, that is, as a collection of coefficients which describe the degree of measured understanding acquired during a learning process from a person B, the expert. The example that was used in the section " a numerical example" shows that within the described formalism, the coefficients can be taken to have fuzzy set values. This character arises out of the way in which these coefficients are assigned their numerical values (despite the fact that the concepts which these coefficients represent may themselves be crisp and well defined). It was also stated as one of the properties of these coefficients that their values had the range $0 \le a_{ij} \le 1.0$, property 5. In addition this was also described in the initial proposition of the concept matrix and the fuzzy character was described in figure 1.1. Hence it would appear that a formal use of fuzzy measures ought to be made in dealing with the process of determining the numerical values of the coefficients.

Description and choices of measures.

Common measures of fuzziness used in the literature, referred to as indices of fuzziness, are defined in terms of metric distances of a fuzzy subset A from any members of the nearest crisp set c. These are the Hamming, and the Minkowsky metrics, of which the Euclidean is a special case of the Minkowsky metric, Klir (1988). The Euclidean and Minkowsky metrics were mentioned earlier in the section under "Mathematical Concept" and written as equations 3.1 and 3.2. The Minkowsky metric, equation 3.2,

$$f(A)=\left(\sum_{x\in Y}[\mu_A(x)-\mu_c(x)]^w\right)^{\frac{1}{w}}$$

is a general metric expression from which it is possible to express the Hamming and Euclidean metrics by assigning the value of w = 1 and w = $\frac{1}{2}$ respectively . $f(A)$ representing the index of fuzziness. On closer inspection it appears that these metrics are in fact the δ_{ij}^k coefficients defined earlier, in particular the Hamming metric. A comparison of each term gives $\mu_A(x)=q_{ij}^k$ and $\mu_c(x)=E_{ij}^k$. However in order to obtain the a_{ij} coefficients an averaging operations was imposed giving equation 5.4. It can be argued that the general Minkowsky metric is also some kind of an aggregate operation and may very well be a better index of fuzziness. It is therefore worth while to consider it in the context of this formalism and subsequently make a comparison. The underlying reason being that since the matrix represents a set of concepts and its corresponding numerical values represent the measured understanding of these concepts, then it is important that these numerical values are as true as it is possible to generate.

The Hamming and Euclidean metrics.

Consider the Hamming metric, from equation 3.1 given by,

$$f(A) = \sum_{x \in X} |\mu_A(x) - \mu_c(x)|$$

and the Euclidean metric, which is the Minkowsky metric for which w = 2 as was written in equation 3.2 but now written as,

$$f(A) = \left(\sum_{x \in X} [\mu_A(x) - \mu_c(x)]^2 \right)^{\frac{1}{2}}$$

in the context of the matrix the above can be written as,

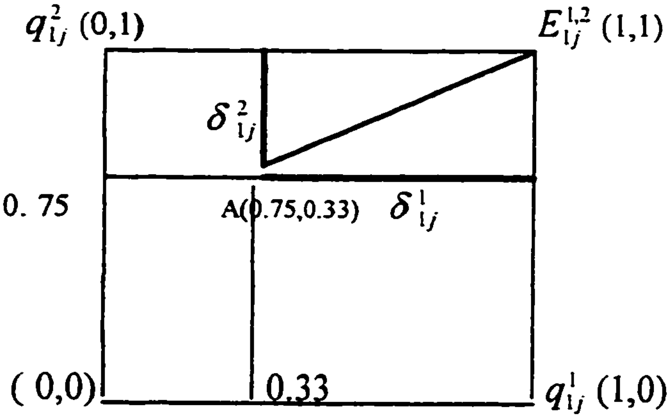
$$f(A) = \sum_k [E_{ij}^k - q_{ij}^k] = \sum_k \delta_{ij}^k \qquad \ldots\ldots\ldots\ldots\ldots\ldots 7.5$$

and $$f(A) = \left(\sum_k [E_{ij}^k - q_{ij}^k]^2 \right)^{\frac{1}{2}} = \left(\sum_k [\delta_{ij}^k]^2 \right)^{\frac{1}{2}} \qquad \ldots\ldots\ldots\ldots 7.6$$

In order to interpret the above equations one must appeal to their geometric aspect. This geometric representation was used earlier, in the section "Proposed Concept Matrix" and shown as a two dimensional case, for two concepts a_{11} and a_{22}. Applying this same diagram in the context of the δ_{ij}^k coefficients then for a two dimensional case one can describe it as shown in figure 7.3.

Figure 7.3

Two dimensional case.



The above geometric representation would be valid if all the E^k_{ij} coefficients were normalized to the value 1.0. Thus if there are n questions posed by the expert relating to the i = 1 and j th concept, then these n questions would form an n dimensional hypercube. For a two dimensional case then the degree of understanding over a domain of knowledge given by the two questions q^1_{1j} and q^2_{1j}, as conversed with the expert, could be defined as the geometric distance of the point A to the vertex (1,1). For this two dimensional case the Euclidean distance is then given by,

$$\delta_{1j}=\left[\left(\delta^1_{1j}\right)^2+\left(\delta^2_{1j}\right)^2\right]^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 7.7$$

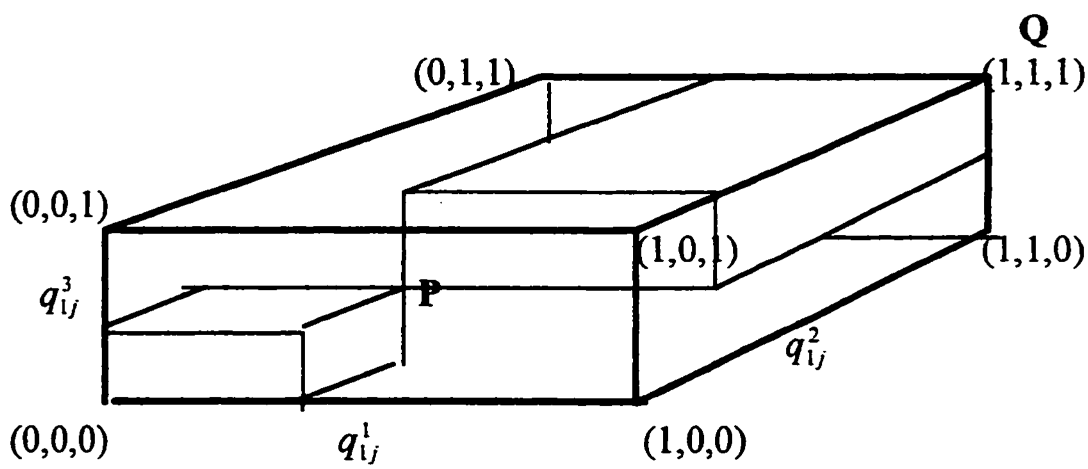
This can be extended to a three dimensional Euclidean space, or higher, quite easily, where the metric distance would be given by,

$$\delta_{1j}=\left[\left(\delta^1_{1j}\right)^2+\left(\delta^2_{1j}\right)^2+\left(\delta^3_{1j}\right)^2\right]^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots 7.8$$

The corresponding geometric diagram for a three dimensional case is shown in figure 7.4 and δ_{ij} is represented by the vector PQ.

Figure 7.4

Three dimensional case.



$$\left(\delta^1_{1j}\right)^2 = (0.6)^2 \quad \left(\delta^2_{1j}\right)^2 = (0.5)^2 \quad \left(\delta^3_{1j}\right)^2 = (0.8)^2$$

In order to compute the numerical values of the a_{ij} coefficients it is necessary to establish the valid algebraic relationship between a_{ij} and the metric coefficient δ_{ij}. This relationship will differ depending on which metric one chooses. The properties of the a_{ij} coefficients will impose the constraints on those relationships. For the Hamming metric, now symbolized as $\{\delta_{ij}\}_H$, equation 5.4 is valid. If one writes $\{\delta_{ij}\}_H = \frac{1}{n}\sum_k \delta^k_{1j}$ then the relationship with the a_{ij} coefficient can be written as,

$$a_{ij} = 1 - \{\delta_{ij}\}_H \qquad \text{7.9}$$

However, for the case involving the Euclidean metric, symbolized as $\{\delta_{ij}\}_E$, the above equation does not satisfy the properties of a_{ij}. This being due to the fact that it is quite possible that $\{\delta_{ij}\}_E$ may acquire a value greater than 1.0

but one cannot determine an average value because $\{\delta_{ij}\}_E$ is unique. A simple 3-D example illustrates this point. Taking the coordinates of point P in figure 16 to be P(0.4,0.5,0.2) then , $\delta_{1j} = \sqrt{0.36 + 0.25 + 0.64} = 1.11$

if one uses equation 49 this will yield the value $a_{ij} = -0.11$ which is in violation of its definition. It is necessary to have a term which when subtracted from the value 1.0 will yield a value which does not exceed 1.0.

To ensure that this does not occur one can consider the ratio $\dfrac{\{\delta_{ij}\}_E}{\{\delta_{ij}\}_{E\max}}$

where $\{\delta_{ij}\}_{E\max}$ is the vector from the point (0,0,...,0) to the vertex point (1,1,....,1). For an n dimensional hypercube $\{\delta_{ij}\}_{E\max} = \sqrt{n}$.

since $\{\delta_{ij}\}_{\max} = \sqrt{= 1^2 + 1^2 + 1^2 + .\ .\ .\ . + 1^2} = \sqrt{n}$ One can thus write for the Euclidean metric,

$$a_{ij} = 1 - \frac{\{\delta_{ij}\}_E}{\sqrt{n}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 7.10$$

where $\{\delta_{ij}\}_E = \left[\sum_k \left(\delta_{ij}^k\right)^2\right]^{\frac{1}{2}}$ and $\delta_{ij}^k = E_{ij}^k - q_{ij}^k$ as defined earlier

Comparative values.

Using the set of questions described in case (i) and assigning arbitrary numerical values for the responses obtained, for i = 1 and j = 1, 2, 3.

the values of a_{ij} can be calculated using equations 6.8 and 6.10.

Table 5 shows the data and calculated values for the δ_{ij} coefficients.

Table 5.
Numerical Values for Q_{ij} and δ_{ij}.

Q_{12}	δ_{12}	$(\delta_{12})^2$	Q_{13}	δ_{13}	$(\delta_{13})^2$	Q_{14}	δ_{14}	$(\delta_{14})^2$
0.60	0.40	0.16	0.52	0.48	0.23	0.25	0.75	0.56
0.45	0.55	0.30	0.44	0.56	0.31	0.46	0.54	0.29
0.32	0.68	0.46	0.81	0.19	0.03	0.65	0.35	0.12
0.74	0.26	0.06				0.55	0.45	0.20
0.25	0.75	0.56						
0.56	0.46	0.21						

From the data in the above table and using equations 7.9 and 7.10 to determine the a_{1j} coefficient, one obtains the following values from the Hamming and Euclidean metrics:

From the Hamming metric : $a_{12} = 0.49$ $a_{13} = 0.49$ $a_{14} = 0.48$
normalized coefficients : $\alpha_{12} = 0.16$ $\alpha_{13} = 0.16$ $\alpha_{14} = 0.16$
hence $U_H = 0.48$

From the Euclidean metric : $a_{12} = 0.53$ $a_{13} = 0.43$ $a_{14} = 0.54$
normalized coefficients : $\alpha_{12} = 0.17$ $\alpha_{12} = 0.14$ $\alpha_{12} = 0.18$
and $U_E = 0.49$

From the results obtained it appears that these are not significantly different. The extent to which these metrics differ depends on the relationship between them. However, the functional relationship is not a simple one and follows some complicated polynomial in n. But it suffices that a difference exists to merit a closer look at the choice of metric.

In considering the criterion to be used in making the choice it would seem that the error factor associated with these procedures ought to be taken into account. A simple error analysis reveals that for a set of errors of δ_{ij}, the resulting errors for $(a_{ij})_E$ is considerably less than that obtained for $(a_{ij})_E$. Therefore on these grounds one would be justified in choosing the Euclidean metric for δ_{ij}.

There is however a more fundamental reason for choosing the Euclidean metric, and that stems from the basic notion that measurements understanding can make sense only in a framework of comparison with respect to another source. Hence a students understanding is measured with respect to that of the teacher or the expert, for example in a tutorial conversation.

<u>Alternate operational definitions for U.</u>

The operational definition for the "understanding " coefficient U was obtained through the defined properties 12, 19 and 21 and was expressed by equation 4.5 as, $U_\alpha = \sum_i \sum_j \alpha_{ij}$ with the normalized coefficients α_{ij} expressed numerically on the line interval [0,1].

As was discussed in the previous section, it is possible to adopt the notion of metrics such as the Hamming and Euclidean in order to gauge the distance or gap between the understanding of the student or learner, A, with respect to the expert, or person B. It seems then reasonable that one could adopt the use of these metrics in the final evaluations of the learning matrix M_A of the student. This would also preserve an internal consistency with respect to the manner in which the coefficients of the matrices are obtained.

Since the properties of the coefficients in the matrix are defined such that

$a_{ij} \leq 1$, property 5. It is possible to replace these coefficients with those representing the difference between α_{ij} and β_{ij} that is $\delta_{ij} = \beta_{ij} - \alpha_{ij}$. One can then obtain a matrix ΔM similar to that of equation 5.1, and at a given time T, written as,

$$\Delta M = \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} & . & . & \delta_{1q} \\ \delta_{22} & \delta_{21} & \delta_{23} & . & . & \delta_{2q} \\ \delta_{33} & \delta_{31} & \delta_{32} & . & . & \delta_{3q} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \delta_{pp} & \delta_{p1} & \delta_{p2} & . & . & \delta_{pq} \end{pmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots 7.11$$

The above matrix describes the difference in knowledge, at a given time. It is thus possible to determine the understanding coefficient U using the notions of the metrics as described by equations 3.1 and 3.2. For the Hamming and Euclidean metrics one has respectively,

$$U_H = 1 - \frac{1}{n}\sum_{i,j}\delta_{ij} \quad \text{and} \quad U_E = 1 - \frac{1}{\sqrt{n}}\left[\sum_{i,j}(\delta_{ij})^2\right]^{\frac{1}{2}}$$

Table 6 displays data taken from table 5 and subsequent results obtained for the three definitions of the understanding coefficients, U_α, U_H and U_E.

Table 6.

Comparison of the U Coefficients.

LEARNING MATRIX						NORMALIZED LEARNING MATRIX					L
0.850	0.900	0.720	0.750	0.640		0.037	0.039	0.031	0.033	0.028	0.17
0.990	0.800	0.760	0.820	0.690		0.043	0.035	0.033	0.036	0.030	0.18
0.980	0.850	0.900	0.870	0.000		0.043	0.037	0.039	0.038	0.000	0.16
0.780	0.950	0.820	0.800	0.770		0.034	0.041	0.036	0.035	0.033	0.18
0.950	0.620	0.710	0.500	0.000		0.041	0.027	0.031	0.022	0.000	0.12
0.150	0.100	0.280	0.250	0.360	1.140	0.023	0.010	0.078	0.063	0.130	0.30
0.010	0.200	0.240	0.180	0.310	0.940	0.000	0.040	0.058	0.032	0.096	0.23
0.020	0.150	0.100	0.130	1.000	1.400	0.000	0.023	0.010	0.017	1.000	1.05
0.220	0.050	0.180	0.200	0.230	0.880	0.048	0.003	0.032	0.040	0.053	0.18
0.050	0.380	0.290	0.500	1.000	2.220	0.003	0.144	0.084	0.250	1.000	1.48

The resulting value for U are : $U_{\alpha} = 0.80$ $U_H = 0.71$ and $U_E = 0.76$

From the above results it is clear that there exists a significant difference between them, which would indicate that a choice between the metrics have to be made. From previous arguments given, it would seem that using the Euclidean metric yields the most reliable value. The same procedure would be applied for the R matrix.

Chapter 8 : Conclusions.

Contribution to the field of Educational Technology.

This discussion relating the contribution that the proposed formalism presented in this thesis makes to the field of educational technology is centered around the definition of what educational technology is and what educational technology aims to achieve.

The definition assumed here is that, educational technology has as its fundamental aim and purpose to facilitate educational and learning processes through the use of technological tools. To do this one needs to consider three aspects of educational or learning processes.

1) A description of the subject matter or body of knowledge that must be learned.
2) The application of a mode of instruction or a teaching modality based on some chosen learning theory.
3) An evaluative process consisting of methods tools and procedures utilized to measure the learning outcome.

Of the three categories described above, the contribution that the formalism proposed in this thesis makes relates to the first and third. The contribution made is in terms of the "tool" that has been chosen, namely the matrix construct. While matrices are common mathematical devices their use in the educational context is not very common.

The proposed formalism demonstrates how, through the use of matrices, one is able to describe and present a given body of knowledge in terms of the concepts or items that make up its content, in a relatively simple and systematic manner. A given discipline can be analyzed as a system of concepts or procedures, and a topic labeled for a cognitive system can also function as a label of performance as well.

Through this formalism it is also shown how the matrices representing the learner and teacher interact to yield a learning outcome. Some general description incorporating the learner's characteristics, types, and the variety of instruction that can be used is given. It is shown that the formalism can accommodate these features if one chooses to engage these levels of complexities.

The advantage that a matrix representation has is inherit in the ability to manipulate the coefficients. This manipulation could constitute an externalization of the interactive process between the learner and the teacher. The precise description of these processes may be quite complex and would , in general, require further and more detailed description within this structure. This aspect of the representation is not dealt with in this thesis. However it could in principle be achieved.

The use of matrices as a tool permits a detailed description of the characteristics of the coefficients that describe it and its operational mode. This in itself is a necessary requirement whenever any tool is utilized. This has been extensively discussed in the Properties and Characteristics of the Coefficient chapter 4.

Since, as was mentioned previously, a matrix construct can be used to describe a system and its coefficients can also function as labels representing performance, this allows for the display of the state of knowledge of a given learner at any given time during the learning process. This matrix technique is facilitated by the fact one can use simple spreadsheet programs to display the coefficients. Once critical values are set for these, it is then possible to make decisions as to what next operation to invoke in the learning process. The numerical values representing the state of knowledge of the learner can be obtained from a built in reservoir of test item questions. This aspect was discussed in chapters 6 and 7.

The additional contribution that this formalism can make is in the evaluative or progress of the learner. Since, in the educational context, one needs to measure the progress of the learner, the model allows for the construction of test item questions within the formalism and the systematic recording of results that will exhibit the knowledge of the learner. It is shown that within this testing structure it is possible to chose different measures. The choices fall within the categories of geometric measures, and, given the simplicity of description and interpretation of the Euclidean metric, it is chosen as the most appropriate. However, the advantage of the proposed formalism is that it can accommodate a wide variety of these measures.

Conclusions and Recommendations for Future Research.

The description of concepts within the construct of matrices have been applied in various fields, in particular when vague concepts are involved, such as query structures, retrieval systems and decision making processes, to name a few. Within the educational sphere, various notions of concept maps have been used extensively, but for most part they have not been formalized into a matrix system. It was shown that the transition from concept mapping to a matrix system representation is a natural one. Such matrix representation immediately lends itself to mathematical operations between its constituent elements, hence the ability to determine detailed needs and possible outcomes. As an example, these matrices are applied to a learning situation between two conversing individuals, a student and teacher or expert, the properties and characteristics of these coefficients in this case have been defined and described. This provides a basis for a description of the full architecture of this knowledge representation, which is the essence of this thesis.

In its defined structure it was shown that it is possible to incorporate the learner's characteristics, or learner type, as well as the model of the expert in terms of mode on instructions chosen. The possibility of applying this representation to real learning situations is an important aspect of the thesis. It was shown that several approaches are available for obtaining numerical values for the measure of learning outcomes. These are obtained primarily through the use of various defined metrics.

A more detailed analysis has shown that it is possible to apply the matrix system representation in situations involving either crisp sets of numerical values or those associated with fuzzy sets. For the fuzzy set cases, one would have to obtain membership curves for the particular learning areas of a given topic. The matrix representation discussed in this thesis has several advantageous features associated with it, as compared with other concept mappings or diagram representations, which are :

a) Since, by definition, this matrix calculus representation necessitates numerical values assigned to the coefficients which represent test results, this affords an ability to make simpler and perhaps more meaningful comparisons when dealing with a set of related concepts.
b) The numerical values for the coefficients can be obtained via conventional testing mechanisms. Since one is interested in the effects that the coefficients have on one another, say representation of the effects of a teaching method on a particular topic or concept, then clearly a numerical representation is called for.
c) The matrix representation enables a wide range of computational approaches, orders and procedures to be adopted. One of its powerful used features is that one need not restricted to use the conventional computational rules of matrix algebra. If this were not the case, it's applicability and capability of generating meaningful results would be severely limited. The whole discipline of teaching and learning is context dependent and as a consequence notations and calculii must allow for a wide range of situations and cases, that this matrix representation can accommodate.

d) By virtue of the numerical association with the coefficients, the matrix system representation is well suited for use in the computer as a computational tool. It greatly facilitates the programming of computational procedures and the speed with which they can be carried out. Thus, resulting outputs can be generated within shorter time intervals. This is important for on-line teaching operations. The use of computers also enables one to deal with matrices of higher dimensions. This in turn allows one to deal with topics that have many complexly related concepts associated with them, define and describe these concepts in greater details.

Further Extension and Research.

One of the primary purpose of this study was to devise a plausible and practicable mathematical matrix representation of concepts associated with topics to be learned or taught in an educational environment.

This study has laid some basic groundwork for this form of knowledge representation, however considerable work and research can and remains to be done with respect to this mathematical framework and its computer implementation. The work to be done lies primarily in two categories, one being the theoretical aspect, the other being experimental.

a) *Theoretical aspect.*

Further research can be done to establish a more rigorous mathematical basis for this formalism. In addition more research will be required to determine further properties, characteristics and mathematical operational procedure associated with the coefficients describing these matrices.

In particular the situations involving fuzzy entities and metaphors, along the number dimensions of association how many and what kind, will require more detailed analysis and defined operational procedures.

b) *Experimental aspect.*

Various applications and examples have been discussed and mentioned in the course of this study. The major application may be for Intelligent Tutoring System, in this regard it would be of interest to apply this proposed construct in a real learning situation and determine its efficacy. It would also be of interest do a comparative study of this matrix construct with other system representations that exist and that are utilized in this field.

Student and teacher measurement modeling are an integral part of this construct, hence more detailed work will be required to establish their interaction and how to operationlize it.

REFERENCES.

Arfken G. (1985). Mathematical Methods for Physicists. Academic Press Inc. 3rd edition. Horcourt Brace Jovanovich Publishers.

Ausubel D. P. (1968). Educational Psychology : A Cognitive View. New York : Holt, Rinehart, & Winston.

Barr A. Feignbaum E.A. (1982). The Handbook of Artificial Intelligence. 2. Los Altos : Kaufmann.

Behnke H. Bachmann F. Fladt K. Suss W. (1986). Fundamental of Mathematics, Volume 1. The MIT press.

Boyd. G. Mitchell P. D. (1992). How can Intelligent CAL better adapt to Learners? Computers and Education Journal, 18. No 1-3, 23-28.

Burus H.L. Capps C.G. (1988). Foundations of Intelligent Tutoring Systems : An Introduction. in M.C. Polson and J.J. Richardson, London : Lawrence, Erlbaum 1-19

Brachman R. J. (1985). On the Epistemological status of Semantic Networks. In Readings in Knowledge Representations. Chapter 10, 191-215. Morgan Kaufmann, Los Altos.

Chen S. M. (1988). A New Approach to Handling Fuzzy Decision Making problems. IEEE Trans. Syst. Man Cyber, 8,ns, 1012-1016.

Dumslaff U. Ebert J. (1992). Structuring the Subject Matter. Computer Assisted Learning. In Tomeck (ed), 4th Intl. Conf ICCAL 1992 Proceedings.

Falmagne J-C. Doignon J-P. Koppen M. Villano M. Johannesen (1990). Introduction to Knowledge Spaces : How to build and search them, Psychological Review, 97, 201-224.

Flagg B.M. (1990). Formative Evaluation for Educational Technology. Hillsdale, NJ : Lawrence Erlbaum Associates.

Fowles G. R. (1977). Analytical Mechanics. Third edition, Holt, Rinehart and Winston.

Frasson C. Ramazani D. (1992). Prediction du Niveau d'Acquisition des Connaissances dans la Modelisation de l'Etudiant. Proceedings of the Second International Conference on Intelligent Tutoring Systems, Montreal. 435-442

Goldstein H. (1965). Classical Mechanics. Addison-Wesley publishing company Inc..

Goldstein, I. P. (1979). The Genetic Graph : A representation for the Evolution of Procedural Knowledge, International Journal of Man-Machine Studies 11, 51-77.

Greeno, J. G. (1976). Psychological Representation of Structured Knowledge. In Structural Learning. Issues and Approaches, 73-95. Joseph M. Scandura. Gordon and Breach Science Publisher.

Gutstein E. (1992). Using Expert Tutor Knowledge to Design a Self-Improving Intelligent Tutoring System. Proceedings of the Second International Conference on Intelligent Tutoring Systems, Montreal. 625-632.

Her G. T Ke J. S. (1983). A Fuzzy Information Retreival System Model. Proc National Computer Symp, 147-155.

Kamel M. Hadfield B. Ismail M. (1990). A Fuzzy Query Processing using Clustering Techniques. Information Processing and Management. 26, n2, 279-293.

Kanvera P. (1988). Sparse Distributed Memory. MIT Press.

Klir G. Folger T. A. (1988). Fuzzy Sets, Uncertainty and Information, Prentice Hall.

Kosko B. (1990). Fuzzy Cognitive Maps. International Journal of Man Machine Studies.

Kosko B. (1992). Neural Networks and Fuzzy Systems. A Dynamical Systems Approach to Machine Intelligence. Prentice Hall. N. J.

Kosko B. (1993). Fuzzy Thinking. The New Science of Fuzzy Logic, Hyperon, Ist ed, New York.

Lipson J. (1967). The Development of Mathematical Curriculum for Individualized Instruction. In J. M. Scandura (ed) Research in Mathematics Education.

Mandl H. Lesgold A. (1988). Learning Issues for Intelligent Tutoring Systems, London Academic Press, 51-77

Marton F. Saljo R. (1976). On Quantitative differences in Learning I: Outcome and Process. British Journal of Education Psychology, 46, 115-127.

Miller A. I. (1987). Imagery in Scientific Thought. The MIT press.

Minsky M.L. Papert S. A. (1990). Perceptrons. 4th ed. The MIT Press.

Mitchell P. D. (1989). Evaluation of understanding in Intelligent Interctive Systems : An application of Conversation Theory. Journal of Interctive Instruction Development, No.3, 7-12

Mitchell P. D. Grogono P. D. (1993). Modelling Techniques for Tutoring Systems. Computers and Education. 20, No.1, 55-61.

Moreira L. N. (1979). Concept maps as a tool for teaching. Journal of College Science Teaching, 8, 283-286

PASK. G. (1975). "The Cybernetics of Human Learning and Performance". Hutchison & Co.

Rindler W. (1977). Essential Relativity, Spring-Verlag.

Scandura J. M. (1976). Deterministic Theorizing in Structural Learning : Three levels of Empiricism. In Structural Learning. Issues and Approaches, J. M. Scandura. Gordon and Breach Science Publisher. 251-273.

Schmid R. Telaro G. (1990). Concept Mapping as an Instructional Strategy for High School Biology. Journal of Educational Research, No.8, 78- 85.

Styblinski M. A. Meyer B. D. (1988). Fuzzy Cognitive Maps. Signal Flow Graphs and Quantitative Cicuit Analysis. Proceedings of the IEEE Intl. Conf. on Neural Networks. vol II 549-556.

Taber W. R. Siegal M (1987). Estimation of Expert Weights with Fuzzy Cognitive Maps. Proceeding of the IEEE Intl. Conf. on Neural Networks (ICNN-87) vol II 319-325.

Taber W. R. (1991). Knowledge Processing with Fuzzy Cognitive Maps. Expert Systems with Application. Vol 2, No.1 83-87.

Taylor S.G. (1990). Concept Mapping in Structural Communication. Unpublished M.A. Thesis. Concordia University.

Terano T. Asai K. Sugeno M. (1992). Fuzzy Systems. Theory and it's Application, Academic Press, Harcourt Brace Jovanovich.

Villano G. (1992). Probabilstic Student Models : Baysian Belief Networks and Knowledge Space Theory. Proceedings of the Second International Conference on Intelligent Tutoring Systems, Montreal, 491-498.

Washmuth. (1988). Situation specific and Situation non-specific Knowledge. In Mandl H. Lesgold A. eds. Learning Isuues for Intelligent Tutoring Systems.

Zadeh L. A. (1965). Fuzzy Sets. Information & Control, No.8, 338-353.

APPENDIX A

Numerical determination of the weight factors.

The weight factors w_{ij} associated with each of the coefficients in the matrix can be determined :

a) Through judgement on the relative importance of each of the coefficients.

b) The judgements are quantified to an extent which permits a quantitative interpretation of these judgements among all the coefficients.

Let $C_1, C_2, C_3, ...C_n$ be the set of items. Quantified judgements on pairs of items, C_i, C_j, are represented by a $p \times q$ matrix, M_A.

Let the comparative ratio factor w_{ij} be defined by $w_{ij} = \frac{\omega_i}{\omega_j}$. Then ω_r represents the rth weight factor associated with the a_tth coefficient.

It is possible to construct an associated matrix of the weight factor ratios, as,

$$W = \begin{pmatrix} w_{11} & w_{12} & . & . & w_{1q} \\ w_{21} & w_{22} & . & . & w_{2q} \\ . & . & . & . & . \\ . & . & . & . & . \\ w_{p1} & w_{p2} & . & . & w_{pq} \end{pmatrix}$$

If one multiplies the above matrix by the column matrix for ω, one has,

$$\begin{pmatrix} w_{11} & w_{12} & . & . & w_{1q} \\ w_{21} & w_{22} & . & . & w_{2q} \\ . & . & . & . & . \\ . & . & . & . & . \\ w_{p1} & w_{p2} & . & . & w_{pq} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ . \\ . \\ \omega_p \end{pmatrix} = \gamma \begin{pmatrix} \omega_1 \\ \omega_2 \\ . \\ . \\ \omega_p \end{pmatrix}$$

One can write the matrix equation, as an eigenvector equation. Following procedures as outlined by Arfken G. (1985) for solving eigenvector equations in the example given, we have,

$$W\omega = \gamma\omega$$
$$(W - \gamma I)\omega = 0$$
$$(W - \gamma)\omega = 0$$

where ω is the eigenvector and γ is the eigenvalue.

An example.

Suppose the matrix contains three coefficients, then there would be three weight factors to be determined. The eigenvector equation. in matrix form will be,

$$\begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & _{23} \\ w_{31} & w_{32} & w_{33} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \gamma \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

We thus have the equation,

$$(W - \gamma)w = 0$$

The only known values are the elements of the matrix W, which are $w_{ij} = \frac{\omega_i}{\omega_j}$

To find the values of w we need to first solve for the eigenvalues. γ. This can be done by setting det $W = 0$, and solving for γ in the conventional manner.

$$\begin{vmatrix} w_{11} - \gamma & w_{12} & w_{13} \\ w_{21} & w_{22} - \gamma & w_{23} \\ w_{31} & w_{32} & w_{33} - \gamma \end{vmatrix} = 0$$

Having obtained the value for the eigenvalue. it is now possible to sove for the eigenvectors. since in the matrix form we have,

$$\begin{pmatrix} w_{11} - \gamma & w_{12} & w_{13} \\ w_{21} & w_{22} - \gamma & w_{23} \\ w_{31} & w_{32} - \gamma & w_{33} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = 0$$

This leads to three simultaneous equations for which the numerical values of $\omega_1, \omega_2, \omega_3$ can be obtained.

The set of simultaneous equations are :

$$
\begin{aligned}
(w_{11} - \gamma)\omega_1 + w_{12}\omega_2 + w_{13}\omega_3 &= 0 \\
w_{21}\omega_1 + (w_{22} - \gamma)\omega_2 + w_{23}\omega_3 &= 0 \\
w_{31}\omega_1 + w_{32}\omega_2 + (w_{33} - \gamma)\omega_3 &= 0
\end{aligned}
$$

Clearly the greater the number of coefficients the the greater the corresponding number of weight factors associated with them, hence correspondigly a larger number of equations to solve.

APPENDIX B.

<u>A Computational example for the determination of</u> ω_i.

Suppose it is desired to obtain the weight factors for the coefficients relating to the study of linear kinematics, discussed in chapter 6. These coefficients can be grouped as shown in the table below.

coefficient label	verbal description	weight factor
a_{12}	Vector representation	ω_1
a_{13}	Linear displacement	ω_2
a_{14}	Linear velocity	ω_3
a_{15}	Linear acceleration	ω_4
a_{16}	Graphical representation	ω_5
a_{17}	Set of equations	ω_6

The associated matrix for the pairing of judgments is the written as,

$$W = \begin{pmatrix} 1 & \frac{\omega_1}{\omega_2} & \frac{\omega_1}{\omega_3} & \frac{\omega_1}{\omega_4} & \frac{\omega_1}{\omega_5} & \frac{\omega_1}{\omega_6} \\ \frac{\omega_2}{\omega_1} & 1 & \frac{\omega_2}{\omega_3} & \frac{\omega_2}{\omega_4} & \frac{\omega_2}{\omega_5} & \frac{\omega_2}{\omega_6} \\ \frac{\omega_3}{\omega_1} & \frac{\varpi_3}{\omega_2} & 1 & \frac{\omega_3}{\omega_4} & \frac{\omega_3}{\omega_5} & \frac{\omega_3}{\omega_6} \\ \frac{\omega_4}{\omega_1} & \frac{\omega_4}{\omega_2} & \frac{\omega_4}{\omega_3} & 1 & \frac{\omega_4}{\omega_5} & \frac{\omega_4}{\omega_6} \\ \frac{\omega_5}{\omega_1} & \frac{\omega_5}{\omega_2} & \frac{\omega_5}{\omega_3} & \frac{\omega_5}{\omega_4} & 1 & \frac{\omega_5}{\omega_6} \\ \frac{\omega_6}{\omega_1} & \frac{\omega_6}{\omega_2} & \frac{\omega_6}{\omega_3} & \frac{\omega_6}{\omega_4} & \frac{\omega_6}{\omega_5} & 1 \end{pmatrix}$$

Since $w_{ij} = \frac{1}{w_{ji}}$, where $w_{ij} = \frac{\omega_i}{\omega_j}$, it is only necessary to obtain judgments for

$\frac{n(n-1)}{2}$ coefficients. Since n = 36 then one needs to obtain judgments for only 15 of the coefficients.

In order to remove any bias in the selection and order in which the pairings are presented to the experts, a random choice of the 15 coefficients is made. This choice is shown in the circled coefficients of the matrix shown.

$$\begin{pmatrix} 1 & \frac{\omega_1}{\omega_2} & \frac{\omega_1}{\omega_3} & \frac{\omega_1}{\omega_4} & \frac{\omega_1}{\omega_5} & \frac{\omega_1}{\omega_6} \\ \frac{\omega_2}{\omega_1} & 1 & \frac{\omega_2}{\omega_3} & \frac{\omega_2}{\omega_4} & \frac{\omega_2}{\omega_5} & \frac{\omega_2}{\omega_6} \\ \frac{\omega_3}{\omega_1} & \frac{\varpi_3}{\omega_2} & 1 & \frac{\omega_3}{\omega_4} & \frac{\omega_3}{\omega_5} & \frac{\omega_3}{\omega_6} \\ \frac{\omega_4}{\omega_1} & \frac{\omega_4}{\omega_2} & \frac{\omega_4}{\omega_3} & 1 & \frac{\omega_4}{\omega_5} & \frac{\omega_4}{\omega_6} \\ \frac{\omega_5}{\omega_1} & \frac{\omega_5}{\omega_2} & \frac{\omega_5}{\omega_3} & \frac{\omega_5}{\omega_4} & 1 & \frac{\omega_5}{\omega_6} \\ \frac{\omega_6}{\omega_1} & \frac{\omega_6}{\omega_2} & \frac{\omega_6}{\omega_3} & \frac{\omega_6}{\omega_4} & \frac{\omega_6}{\omega_5} & 1 \end{pmatrix}$$

A verbal scale is devised in which judgments of the " relative importance " is required. In the survey presented, the respondants only need make a selection from the scale shown.

Note: The numeric scale does not appear in the form given to the experts.

	5	3	1	3	5	
"L" left item	"L much more important than "R"	"L" more important than "R"	both equal	"R" more important than "L"	"R" much more important than "L"	"R" right item

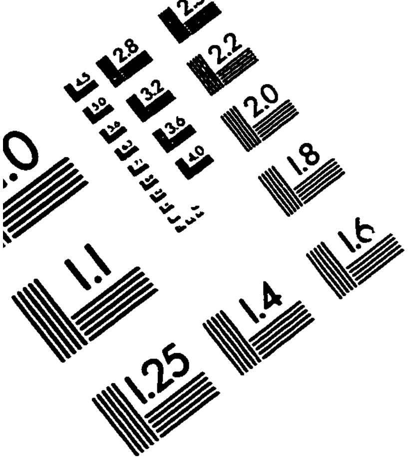
The table below would be the actual form given submitted to the experts to fill, based on their personal judgement and expertise, from the circled elements in the matrix described previously.

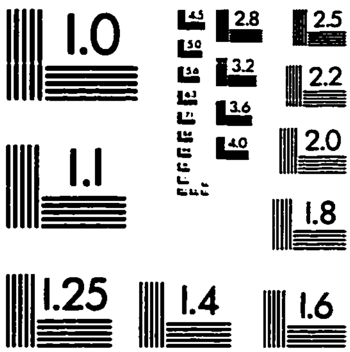
"L" left item	"L" much more important than "R"	"L" more important than "R"	both equal	"R" more important than "L"	"R" much more important than "L"	"R" right item
velocity						graphs
equations						velocity
graphs						displacement
displacement						vectors
acceleration						equations
vectors						graphs
graphs						acceleration
equations						velocity
acceleration						vectors
displacement						acceleration
vectors						velocity
graphs						equations
velocity						displacement
displacement						equations
velocity						acceleration

List of Figures.

List of Tables.

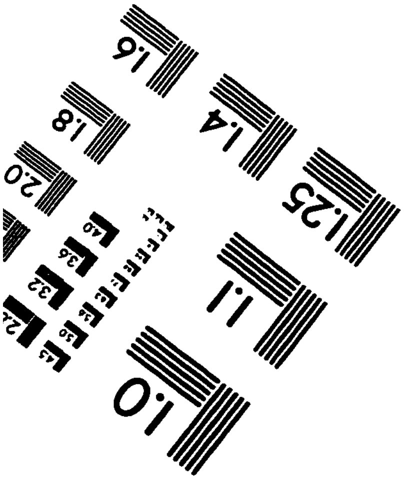
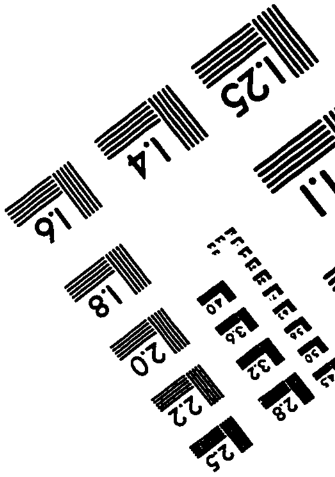
TEST TARGET (QA-3)





150mm

6″

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989