RR-91-49-ONR

ITEM CONSTRUCTION AND PSYCHOMETRIC MODELS APPROPRIATE FOR CONSTRUCTED RESPONSES

Kikumi K.Tatsuoka

This research was sponsored in part by the
Cognitive Science Program
Cognitive and Neural Sciences Division
Office of Naval Research, under
Contract No. N00014-90-J-1307
R&T 4421559

Kikumi K. Tatsuoka, Principal Investigator

Educational Testing Service
Princeton, New Jersey

August 1991

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION
Unclassified

1b RESTRICTIVE MARKINGS

2a SECURITY CLASSIFICATION AUTHORITY

2b DECLASSIFICATION / DOWNGRADING SCHEDULE

3 DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S)
RR-91-49-ONR

5 MONITORING ORGANIZATION REPORT NUMBER(S)

6a NAME OF PERFORMING ORGANIZATION
Educational Testing Service

6b OFFICE SYMBOL (If applicable)

7a NAME OF MONITORING ORGANIZATION
Cognitive Science Program, Office of Naval Research (Code 1142CS)

6c ADDRESS (City, State, and ZIP Code)
Princeton, New Jersey 08541

7b ADDRESS (City, State, and ZIP Code)
800 N. Quincy Street
Arlington, VA 22217-5000

8a NAME OF FUNDING / SPONSORING ORGANIZATION

8b OFFICE SYMBOL (If applicable)

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-90-J-1307

8c ADDRESS (City, State, and ZIP Code)

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153 N	RR 04204	RR04204-01	R&T4421559

11 TITLE (Include Security Classification)
Item construction and psychometric models appropriate for constructed responses (unclassified)

12 PERSONAL AUTHOR(S)
Kikumi K. Tatsuoka

13a TYPE OF REPORT
Technical

13b TIME COVERED
FROM 1989 TO 1992

14 DATE OF REPORT (Year, Month, Day)
August 1991

15 PAGE COUNT
50

16 SUPPLEMENTARY NOTATION

17 COSATI CODES

FIELD	GROUP	SUB-GROUP
05	10	

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
item construction; cognitive structure; adult literacy; graph theory; cognitive processes; Boolean algebra; fraction addition

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

ABSTRACT

Constructed-response formats are desired for measuring complex and dynamic response processes which require the examinee to understand the structures of problems and micro-level cognitive tasks. These micro-level tasks and their organized structures are usually unobservable. This study shows that

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a NAME OF RESPONSIBLE INDIVIDUAL
Dr. Susan Chipman

22b TELEPHONE (Include Area Code)
703-696--318

22c OFFICE SYMBOL
ONR 1142CS

DD Form 1473, JUN 86 — Previous editions are obsolete — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Unclassified

elementary graph theory is useful for organizing these micro-level tasks and for exploring their properties and relations. Moreover, this approach enables us to better understand macro-level performances on test items. Then, an attempt to develop a general theory of item construction is described briefly and illustrated with the domains of fraction addition problems and adult literacy. Psychometric models appropriate for various scoring rubrics are discussed.

Item Construction and Psychometric Models Appropriate for Constructed Responses

Kikumi K. Tatsuoka

August 1991

ABSTRACT

Constructed-response formats are desired for measuring complex and dynamic response processes which require the examinee to understand the structures of problems and micro-level cognitive tasks. These micro-level tasks and their organized structures are usually unobservable. This study shows that elementary graph theory is useful for organizing these micro-level tasks and for exploring their properties and relations. Moreover, this approach enables us to better understand macro-level performances on test items. Then, an attempt to develop a general theory of item construction is described briefly and illustrated with the domains of fraction addition problems and adult literacy. Psychometric models appropriate for various scoring rubrics are discussed.

Introduction

Recent developments in cognitive theory suggest that new achievement tests must reflect four important aspects of performance: The first is to assess the principle of performance on a test that is designed to measure, the second is to measure dynamic changes in students' strategies, the third is to evaluate the structure or representation of knowledge and cognitive skills, and the fourth is to assess the automaticity of performance skills (Graser, 1985).

These measurement objectives require a new test theory that is both qualitative and quantitative in nature. Achievement measures must be both descriptive and interpretable in terms of the processes that determine performance. Traditional test theories have shown a long history of contributions to American education through supporting norm-referenced and criterion-referenced testing.

Scaling of test scores has been an important goal in these types of testing, while individualized information such as diagnosis of misconceptions has never been a main concern of testing. In these contexts the information objectives for a test will depend on the intended use of the test. Standardized test scores are useful for admission or selection purposes but such scores cannot provide teachers with useful information for designing remediation. Formative uses of assessment require new techniques, and this chapter will try to introduce one of such techniques.

Constructed-response formats are desirable for measuring complex and dynamic cognitive processes (Bennett, Ward, Rock, & LaHart, 1990) while multiple-choice items are suitable for measuring static knowledge. Birenbaum and Tatsuoka (1987) examined the effect of the response format on the diagnosis of examinees' misconceptions and concluded that multiple-choice items may not provide appropriate information for identifying students' misconceptions. The constructed-response format, on the other hand, appears to be more appropriate. This finding also confirms the assertion mentioned above by Bennett et al. (1990).

As for the second objective, several studies on "bug" stability suggest that bugs tend to change with "environmental challenges" (Ginzburg, 1977) or "impasses" (Brown & VanLehn, 1980). Sleeman and his associates (1989) developed an intelligent tutoring system aimed at the diagnosis of bugs and their remediation in algebra. However, bug instability made diagnosis uncertain and hence remediation could not be directed. Tatsuoka, Birenbaum and Arnold (1990) conducted an experimental study to test the stability of bugs and also found that inconsistent rule application was common among students who had not mastered signed-number arithmetic operations. By contrast, mastery-level students showed a stable pattern of rule application. These studies strongly indicate that the unit of diagnosis should be neither erroneous rules nor bugs but somewhat larger components such as sources of misconceptions or instructionally relevant cognitive components.

The primary weakness of attempts to diagnose bugs is that bugs are tentative solutions for solving the problems when students don't have the right skills.

However, the two identical subtests (32 items each) used in the signed-number study, had almost identical true score curves for the two parameter-logistic model (Tatsuoka & Tatsuoka, 1991). This means that bugs are unstable but total scores are very stable. Therefore, searching for the stable components that are cognitively relevant is an important goal for diagnosis and remediation.

The third objective, evaluating the structure or representation of cognitive skills, requires response formats different from traditional item types. We need items that ask examinees to draw flow charts in which complex relations among tasks, subtasks, skills and solution path are expressed graphically, or that ask examinees to describe such relations verbally. Questions can be figural response formats in which examinees are asked to order the causal relationships among several concepts and connect them by a directed graph.

These demanding measurement objectives apparently require a new psychometric theory that can accommodate more complicated forms of scoring than just right or wrong item-level responses. The correct response to the item is determined by whether or not all the cognitive tasks involved in the item can be answered correctly. Therefore, the hypothesis in this regard would be that if any of the tasks would be wrong, then there would be a high probability that the final answer would also be wrong.

These item-level responses are called macro-level responses and those of the task-level are called micro-level responses. This report will address such issues as follows:

The first section will discuss macro-level analyses versus micro-level analyses and will focus on the skills and knowledge that each task requires.

The second section will introduce elementary graph theory as a tool to organize various micro-level tasks and their directed relations.

Third, a theory for designing constructed-response items will be discussed and will be illustrated with real examples. Further, the connection of this deterministic approach to the probabilistic models, Item Response Theory and Rule space models (Tatsuoka, 1983, 1990) will also be explained. These models will be demonstrated as a computation device for drawing inferences about micro-level performances from the item-level responses.

Finally, possible scoring rubrics suitable for graded, continuous and nominal response models will be addressed.

Macro- And Micro-Level Analyses

Making Inferences On Unobservable Micro-Level Tasks From Observable Item-Level Scores

Statistical test theories deal mostly with test scores and item scores. In this study, these scores are considered to be macro-level information while the underlying cognitive processes are viewed as micro-level information. Here we shall be using a much finer level of observable performances than the item level or the macro-level.

Looking into underlying cognitive processes and speculating about examinees' solution strategies, which are unobservable, may be analogous to the situation that modern physics has come through in the history of its development. Exploring the properties and relations among micro-level objects such as atoms, electrons, neutrons and other elementary particles, has led to many phenomenal successes in theorizing about physical phenomena at the macro-level such as the relation between the loss and gain of heat and temperature. Easley and Tatsuoka (1968) state in their book Scientific Thought that "the heat lost or gained by a sample of any non-atomic substance not undergoing a change of state is jointly proportional to the number of atoms in the sample and to the temperature change. This strongly suggests that both heat and temperature are intimately related to some property of atoms." Heat and temperature relate to molecular motion and the relation can be expressed by mathematical equations involving molecular velocities.

This finding suggests that, analogously, it might be useful to explore the properties and relations among micro-level and invisible tasks, and to predict their outcomes. These are observable as responses to test items. The approach mentioned above is not new in scientific research. In this instance, our aim is to explore a method that can, scientifically, explain macro-level phenomena -- in our context item-level or test-level achievement -- derived from micro-level tasks. The method should be generalizable from specific relations in a specific domain to general relations in general domains. In order to accomplish our goal, elementary graph theory is used.

Identification of Prime Subtasks or Attributes

The development of an intelligent tutoring system or cognitive error diagnostic system, involves a painstaking and detailed task analysis in which goals, subgoals and various solution paths are identified in a procedural network (or a flow chart). This process of uncovering all possible combinations of subtasks at the micro-level is essential for making a tutoring system perform the role of the master teachers, although the current state of research in expert systems only partially achieves this goal. According to Chipman, Davis and Shafto (1986), many studies have shown the tremendous effectiveness of individual tutoring by master teachers.

It is very important that analysis of students' performances on a test be similar to various levels of analyses done by human teachers while individual tutoring is given. Although the context of this discussion is task analysis, the methodology to be introduced can be applied in more general contexts such as skill analysis, job analysis or content analysis.

Identifying subcomponents of tasks in a given problem-solving domain and abstracting their attributes is still an art. It is also necessary that the process be made automatic and objective. However, we here assume that the tasks are already divided into components (subtasks) and that any task in the domain can be expressed by a combination of cognitively relevant prime subcomponents. Let us denote these by $A_1, \ldots, A_k$ and call them a set of attributes.

Insert Figure 1 about here

Determination of Direct Relations Between Attributes

Graph theory is a branch of mathematics that has been widely used in connection with tree diagrams consisting of nodes and arcs. In practical applications of graph theory, nodes represent objects of substantive interest and arcs show the existence of some relationship between two objects. In the task-analysis setting, the objects correspond to attributes. Definition of a direct relation is determined by the researcher using graph theory, on the basis of the purpose of his/her study.

For instance, $A_k \rightarrow A_l$ if A_k is an immediate prerequisite of A_l (Sato, 1990), or $A_k \rightarrow A_l$ if A_k is easier than A_l (Wise,1981). These direct relations are rather logical but there are also studies using sampling statistics such as proximity of two objects (Hubert, 1974) or dominance relations (Takeya, 1981). (See M. Tatsuoka (1986) for a review of various applications of graph theory in educational and behavioral research.)

The direct relations defined above can be represented by a matrix called the adjacency matrix $\mathbf{A} = (a_{kl})$ where

$$\begin{cases} a_{kl} = 1 \text{ if a direct relation exists from } A_k \text{ to } A_l \\ a_{kl} = 0 \text{ otherwise} \end{cases}$$

If a direct relation exists from A_k to A_l and also from A_l to A_k, then A_k and A_l are said to be equivalent. In this case, the elements a_{kl} and a_{lk} of the adjacency matrix are both one.

There are many ways to define a direct relationship between two attributes, but we will use a "prerequisite" relation in this paper. One of the open-ended questions shown in Bennett et al. (1990) will be used as an example to illustrate various new terminologies and concepts in this study.

Item 1: How many minutes will it take to fill a 2,000-cubic-centimeter tank if water flows in at the rate of 20 cubic-centimeters per minute and is pumped out at the rate of 4 cubic-centimeter per minute?

This problem is a two-goal problem and the main canonical solution is that:

1. Net filling rate = 20 cc per minute - 4 cc per minute
2. Net filling rate = 16 cc per minute
3. Time to fill tank = 2000 cc/16 cc per minute
4. Time to fill tank = 125 minute.

Let us define attributes involved in this problem:

A_1 : First goal is to find the net filling rate
A_2 : Compute the rate
A_3 : Second goal is to find the time to fill the tank
A_4 : Compute the time.

In this example, A_1 is a prerequisite of A_2, A_2 is a prerequisite of A_3, and A_3 is a prerequisite of A_4. This relation can be written by a chain, A_1 -> A_2 -> A_3 -> A_4. This chain can be expressed by an adjacency matrix whose cells are

$a_{12} = a_{23} = a_{34} = 1$, and others are zeros.

$$\text{Adjancency matrix} \quad A = \begin{array}{c} \begin{array}{cccc} A_1 & A_2 & A_3 & A_4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array}$$

This adjacency matrix **A** is obtained from the relationships among the attributes which are required for solving item 1. The prerequisite relations expressed in the adjacency matrix **A** in this example may change if we add new items. For instance, if a new item -- that requires only the attributes A_3 and A_4 to reach the solution -- is added to the item pool consisting of only item 1, then A_1 may not be considered as the prerequisite of A_3 any more. The prerequisite relation, in practice, must be determined by a task analysis of a domain and usually it is independent of items that are in an item pool.

<u>Reachability Matrix: Representation of All the Relations, Both Direct and Indirect</u> Warfield (1973a,b) developed a method called "interactive structural modeling" in the context of switching theory.

By his method, the adjacency matrix shown above indicates that there are direct relations from A_1 to A_2, from A_2 to A_3 and from A_3 to A_4 but no direct relations other than among these three arcs. However, a directed graph (or digraph) consisting of A_1, A_2, A_3, and A_4 shows that there is an indirect relation from A_1 to A_3, from A_2 to A_4, and A_1 to A_4.

Warfield showed that we can get a reachability matrix by multiplying the matrix **A** + I -- the sum of the adjacency matrix **A** and the identity matrix I -- by itself n times in terms of Boolean Algebra operations. The reachability matrix indicates that reachability is at most n steps (A_k to A_l), whereas the adjacency matrix contains reachability in exactly one step (A_k to A_l) [a node is reachable from itself in zero steps]. The reachability matrix of the example in the previous section is given below:

$$\mathbf{R} = (\mathbf{A} + \mathrm{I})^3 = (\mathbf{A} + \mathrm{I})^4 = (\mathbf{A} + \mathrm{I})^5 = \ldots.$$

$$\mathbf{R} = \begin{array}{cc} \begin{array}{cccc} A_1 & A_2 & A_3 & A_4 \end{array} & \\ \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \end{array}$$

where the definition of Boolean operations is as follows:

1 + 1 = 1, 1 + 0 = 0 + 1 =1, 0 + 0 = 0 for addition and

1 x 1 = 1, 0 x 1 = 1 x 0 = 0, 0 x 0 = 0 for multiplication.

The reachability matrix indicates that all attributes are related directly or indirectly. From the chain above, it is obvious that although A_k and A_{k+1} relate directly A_k and A_{k+2} relate indirectly.

This form of digraph representation of attributes can be applied to either evaluation of instructional sequences, curriculum evaluation, and documentation analysis and has proved to be very useful (Sato, 1990). Moreover, reachability matrix can provide us with information about cognitive structures of attributes. However, application to assessment analysis requires extension of the original method introduced by Warfield.

A Theory of Item Design Appropriate For The Constructed-Response Format

An Incidence Matrix In Assessment Analysis

The adjacency matrix (a_{kl}) is a square matrix of order K x K, where K is the number of attributes and a_{kl} represents the existence or absence of a direct directed relation from A_k to A_l. Let us consider a special case.

When the adjacency matrix **A** is a null matrix, hence **A** + **I** is the identity matrix of the order k -- there is no direct relation among the attributes. Let Ω be a set $\{A_1, A_2, \ldots, A_k\}$ and L be the set of all subsets of Ω,

$$L = [\{A_1\}, \{A_2\}, \ldots, \{A_1, A_2\}, \{A_1, A_3\}, \ldots, \{A_1, A_2 \ldots, A_k\}, \{\}],$$

then L is called a lattice in which the number of elements in L is 2^k.

In this case, we should be able to construct an item pool of 2^k items in such a manner that each item involves only one element of L. There is a row for each attribute and a column for each item, and the element of 1 in (k,j)-cell indicates that item j involves attribute A_k while 0 indicates that item j does not involve A_k. Then this matrix of order K x 2^k -- or K x n for short -- is called an incidence matrix, $Q = (q_{kj})$, k=1,...K & j=1,...n.

For example, in the matrix Q below, k + 1 th column (item k + 1) has the vector of (1 1 0 ... 0) which corresponds to the k + 1 th set, $\{A_1, A_2\}$ in L.

$$\begin{array}{ll} & \begin{array}{ccccccccc} i1 & i2 & \cdots & ik & i(k+1) & i(k+2) & \cdots & i(2_k-1) & i(2^k) \end{array} \\ Q(kxn) = & \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_k \end{array} \end{array}$$

However, if K becomes large, say K=20, then the number of items in the item pool becomes astronomically large, 2^{20}=1,048,576. In practice, it might be very difficult to develop a pool of constructed response items so that each item requires only one independent attribute. Constructed response items are usually designed to measure such functions as cognitive processes, organization of knowledge and cognitive skills, and theory changes required in solving a problem. These complex mental activities require an understanding of all the relationships which exist in the elements of Ω. Some attributes are connected by a direct relation while others are isolated.

In general, the manner in which the attributes in Ω interrelate, one with another, bear a closer resemblance to the arc/node tree configuration than they do to the unidimensional chain shown in the previous section.

Suppose we modify the original water-filling-a-tank problem to make four new items (beyond our original item 1 - page 8), which include the original attributes.

Item 2 What is the net filling rate of water if water flows in at the rate of 50 cc/min and out at the rate of 35 cc/min ?

Item 3 What is the net filling rate of water if water flows in at the rate of h cc/min and out at the rate of d cc/min?

Item 4 How many minutes will it take to fill a 1,000-cubic-centimeters tank if water flows in at the rate of 50 cubic-centimeters per minutes?

Item 5 How many minutes will it take to fill an x cubic-centimeters water tank if water flows in at the rate of y cubic-centimeters per minutes?

The incidence matrix Q for the five items will be:

$$Q(4x5) = \begin{array}{c} \begin{array}{ccccc} i1 & i2 & i3 & i4 & i5 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array}$$

The prerequisite relations among the four attributes are changed from the "totally ordered" chain, $A_1 \rightarrow A_2 \rightarrow A_3 \rightarrow A_4$ to the partially ordered relation as stated below. That is, A_1 is a prerequisite of A_2, A_3 is a prerequisite of A_4, but A_2 is not a prerequisite of either A_3 or A_4. The relationship among the attributes is no longer a totally-ordered chain but two totally-ordered chains, $A_1 \rightarrow A_2$ and $A_3 \rightarrow A_4$.

Tatsuoka (1991) introduced the inclusion order among the row vectors of an incidence matrix and showed that a set of the row vectors becomes Boolean Algebra with respect to Boolean addition and multiplication. In this Boolean algebra, the prerequisite relation of two attributes becomes equivalent to the inclusion order between two row vectors -- that is, the row vectors A_1 and

A_3 include the row vectors A_2 and A_4, respectively, in the Q(4 x 5) matrix above.

There is an interesting relationship between an incidence matrix Q(k x n) and the reachability matrix R(k x k). A pairwise comparison over all the combinations of the row vectors of Q(k x n) matrix with respect to the inclusion order will yield the reachability matrix R(k x k) in which all the relations logically existing among the k attributes, both direct or indirect, are expressed. This property is very useful for examining the quality and cognitive structures of an item pool.

The adjacency and reachability matrices of the GRE items given earlier are given below:

$$\mathbf{A}(4\text{x}4) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{R}(4\text{x}4) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

However, the reachability matrix of the case given in Q(kxn) in which k attributes have no relations will be the identity matrix of the order k. This result can be easily confirmed by examining the inclusion relation of all pairs of the row vectors of the matrix Q(k x n).

Connection of our Deterministic Approach to Probability Theories

Tatsuoka and Tatsuoka (1987) introduced the slippage random variable S_j, which is assumed to be independent across the items, as follows:

If $S_j = 1$, then $X_j = 1 - R_i$ and if $S_j = 0$, then $X_j = R_j$.

or, equivalently, $S_j = \left| X_j - R_j \right|$.

A set $\{X_m\}$ forms a cluster around R -- (where X_m is an item response pattern that is generated by adding different numbers of slips to the ideal item pattern R). The Tatsuokas showed that the total number of slippage s in these "fuzzy" item patterns follows a compound binomial distribution with the slippage probabilities unique to each item. They called this distribution the "bug distribution."

However, it is also the conditional distribution of s given R, where R is a state of knowledge and capabilities. This is called a state distribution for short. Once a distribution is determined for each state of knowledge and capabilities, then Bayes' decision rule for minimum errors can be applied to classify any student's response patterns into one of these predetermined states of knowledge and capabilities (Tatsuoka & Tatsuoka, 1987).

The notion of classification has an important implication for education. Given a response pattern, we want to determine the state to which the students' misconception is the closest and we want to answer the question: "What misconception, leading to what incorrect rule of operation, did this subject most likely have?" or "What is the probability that the subject's observed responses have been drawn from each of the predetermined states?" This is error diagnosis.

For Bayes' decision rule for minimum errors, the classification boundary of two groups of "fuzzy" response patterns becomes the linear discriminant function when the state distributions are a multivariate normal and their covariance matrices are approximately equal. Kim (1990) examined the effect of violation of the normality requirement, and found that the linear discriminant function is robust against this violation. Kim further compared the classification results using the linear discriminant functions and K nearest neighbors method, which is a non-parametric approach, and found that the linear discriminant functions are better. However, the classification in the n-dimensional space with many predetermined groups (as many as 50 or 100 states) is not practical.

Tatsuoka (1983, 1985, 1990) proposed a model (called 'rule space') that is capable of diagnosing cognitive errors. Rule space uses item response functions where the probability of correct response to item j is modeled as a function of the student's "proficiency", (which is denoted by θ) as $P_j(\theta)$, and that $Q_j(\theta)=1-P_j(\theta)$. Since the rule space model maps all possible item response patterns into ordered pairs of (θ,ζ) and where ζ is an index measuring atypicality of response patterns (a projection operator by a mathematical term), all the error groups will also be mapped into this Cartesian Product space. The mapping is one-to-one at almost everywhere if IRT functions are monotone increasing (Tatsuoka, 1985; Dibello & Baillie, 1991).

Figure 3 illustrates the rule space configuration.

Insert Figure 3 about here

Rule space can be regarded as a technique for reducing the dimensionality of the classification space. Furthermore, since the clusters of "fuzzy" response patterns that are mapped into the two dimensional space follow approximately bivariate normal distributions (represented by the ellipses shown in Figure 3), Bayes' decision rules can be applied to classify a point in the space into which one of the ellipses shown in Figure 3), (M. Tatsuoka & K Tatsuoka, 1989; Tatsuoka, 1990).

Kim also compared the classification results using rule space with Bayes' classifiers -- the discriminant function approach -- and the non-parametric _K_-nearest neighbors method. He found that the rule space approach was efficient in terms of CPU time, and that the classification errors were as small as those created by the other two methods.

Moreover, states located in the two extreme regions of the θ scale, tended to have singular within-groups covariance matrices in the n-dimensional space; hence, classification using discriminant functions could not be carried out for such cases. The rule space classification, on the other hand, was always obtainable and reasonably reliable.

We assumed the states for classification groups were pre-determined. However, determination of the universal set of knowledge states is a complicated task and it requires a mathematical tool, Boolean algebra, to cope with the problem of combinatorial explosion (Tatsuoka, 1991).

We utilized a deterministic logical analysis to narrow down the fuzzy region of classification as much as possible to the extent that we would not lose the interpretability of misconceptions and errors. Then the probability notion, used to explain such uncertainties as instability of human performances on items, was used to express perturbations.

Correspondence Between the Two Spaces, Attribute Responses and Item Responses

Tatsuoka (1991), Varadi & Tatsuoka (1989) introduced a "Boolean descriptive function" f to establish a relationship between the attribute responses and item responses.

For example, in the matrix Q(4 x 5), a subject who can not do A_1 but can do A_2, A_3, and A_4, will have the score of 1 for those items that do not involve A_1 and the score of 0 for those that do involve A_1. Thus, the attribute pattern (0 1 1 1) corresponds to the observable item pattern (0 0 0 1 1).

By making the same kinds of hypothesis on the different elements of L and applying these hypotheses to the row vectors of the incidence matrix Q, we can derive the item patterns that are logically possible for a given Q matrix. These item patterns are called ideal item patterns (denoted by Ys).

Generally speaking, the relationship between the two spaces, the attribute and item spaces is not straightforward as the example of Q(4 x 5). This is because partial order relations among the attributes almost always exist and a given item pool often does not include the universal set of items which involve all possible combinations of attributes.

<u>A case when there is no relation among the attributes</u>

Suppose there are four attributes in a domain of testing, and that the universal set of items 2^4 are constructed, then incidence matrix of 2^4 items is given below:

$$Q(4 \times 16) = \begin{array}{c} \begin{array}{cccccccccccccccc} 1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16 \end{array} \\ \left[\begin{array}{cccccccccccccccc} 0&1&0&0&0&1&1&1&0&0&0&1&1&1&0&1 \\ 0&0&1&0&0&1&0&0&1&1&0&1&1&0&1&1 \\ 0&0&0&1&0&0&1&0&1&0&1&1&0&1&1&1 \\ 0&0&0&0&1&0&0&1&0&1&1&0&1&1&1&1 \end{array}\right] \end{array} \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array}$$

An hypothesis that states "this subject cannot do A_l but can do $A_1, .. A_{l-1}$, $A_{l+1}, .. A_k$ correctly" corresponds to the attribute pattern (1 ...1 0 1...1). Let us denote this attribute pattern by Y_l, then $\mathbf{Y}_l$ produces the item pattern $\mathbf{X}_l$ where $x_j = 1$ if item j does not involve A_l, and $x_j = 0$ if item j involves A_l. This operation is defined as a Boolean descriptive function.

Sixteen possible attribute patterns and the images of f (16 ideal item patterns), are summarized in Table 1 below.

Insert Table 1 about here

For instance, attribute response pattern 1 0 indicates that a subject cannot do A_1 and A_3 correctly but can do A_2 and A_4. Then from the incidence matrix Q(4x16) shown above, we see that the scores of items 2,4,6,7,8,9 11,12,13,14,16 must become zero while the scores of 1,3,5,10 must be 1.

Table 1 indicates that any responses to the 16 items can be classified into one of the 16 predetermined groups. They are the universal set of knowledge and capability states that are derived from the incidence matrix Q(4 x 16) by applying the properties of Boolean algebra. In other words, the 16 ideal item patterns exhaust all the possible patterns logically compatible with the constraints imposed by the incidence matrix Q(4 x 16). By examining and comparing a subject's responses with these 16 ideal item patterns, one can infer the subject's performances on the unobservable attributes. As long as these attributes represent the true task analysis, any response patterns of the above 16 items, which differ from the 16 ideal item patterns, are regarded as fuzzy patterns or perturbations resulting from some lapses or slips on one or more items, reflecting random errors.

A Case When There Are Prerequisite Relations Among the Attributes

So far we have not assumed any relations among the four attributes in Table 1. It is often the case that some attributes are directly related one to another. Suppose A_1 is a prerequisite of A_2, A_2 is a prerequisite of A_3 and A_1 is also a prerequisite of A_4.

Insert Figure 2 about here

If we assume that a subject cannot do A_1 correctly, then A_2 and A_3 cannot be correct because they require knowledge of A_1 as a prerequisite. Therefore, the attribute patterns 3, 4, 5, 9, 10, 11, and 15 in Table 1 become (0 0 0 0) which is pattern 1.

By an argument similar to the above paragraph, "cannot do A_2 " implies "cannot do A_3". In this case the attribute patterns 2 and 7, and the patterns 8 and 14 are respectively no longer distinguishable. Table 2 summarizes the implication of the relations assumed above among the four attribute set.

Insert Table 2 about here

The number of attribute patterns has been reduced from 16 to 7. The item patterns associated with these seven attribute patterns are given in the right-hand column, in which each pattern still has 16 elements. It should be noted that we do not need 16 items to distinguish seven attribute patterns. Items 2, 3, 4, 5, 10, and 11 are sufficient to provide the different ideal item patterns, (0 0 0 0 0 0), (1 0 0 0 0 0 0), (1 0 0 1 0 0), (1 1 0 1 1 0), (1 1 0 0 0 0), (1 1 1 0 0 0), (1 1 1 1 1 1), which are obtained from the second through fifth columns, and the 10th and 11th columns of the ideal item patterns in Table 2.

The seven reduced attribute paterns given in Table 2 can be considered as a matrix of the order 7 x 4. The four column vectors, which associate with attributes, A_1, A_2, A_3 and A_4 satisfy the partial order defined by the inclusion relation. Expressing the inclusion relationships among the four attributes -- A_1 (column 1), A_2 (column 2), A_3 (column 3) and A_4 (column 4) -- in a matrix, results in the following reachability matrix **R**:

$$R = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It is easy to verify that **R** can be derived from the adjacency matrix of **A** obtained from the prerequisite relations among the four attributes; A_1 -> A_2 -> A_3 and A_1 -> A_4.

An approach to design constructed-response items for a diagnostic test.

Notwithstanding the above, it is sometimes impossible to construct items like 2,3,4, and 5 which involve only one attribute per item. This is especially true when we are dealing with constructed-response items, we have to measure much more complicated processes such as organization of knowledge and cognitive tasks. In these cases, it is natural to assume that each item will involve several attributes. By examining Table 2, one can find several sets of items for which the seven attribute patterns produce exactly the same seven ideal item patterns as those in Table 2.

For example, they are a set, {2,3,4,5,10,11}, or {2,3,4,5,13,11}. These two sets of items are just examples which are quickly obtained from Table 2. There are 128 different sets of items which produce the seven ideal item patterns when the seven attribute patterns in Table 2 are applied. This means that there are many possibilities for selecting an appropriate set of six items so as to maximize diagnostic capability of a test. The common condition for selection of these sets of items can be generalized by the use of Boolean algebra, but detailed discussion will not be given in this paper.

This simple example implies that this systematic item construction method enables us to measure unobservable underlying cognitive processes via observable item response patterns. However, if the items are constructed without taking these requirements into account, then instructionally useful feedback or cognitive error diagnoses may not be always obtainable.

Explanation with GRE math items

The five items associated with GRE water filling problem are given in the earlier section. The incidence matrix Q(4 x 5) produces nine ideal item patterns and attribute patterns by using BUGLIB program (Varadi & Tatsuoka, 1989). Table 3 summarizes them.

Insert Table 3 about here

The prerequisite relations, A_1 -> A_2 and A_3 -> A_4 imply some constraints on attribute patterns: the attribute pattern, (0 1) for A_1, A_2 and A_3, A_4 cannot exist logically. A close examination of Table 1 reveals that the constraints result in nine distinguishable attribute patterns. They are: 3,5,10 result in 1 that is (0000); 8 to 2 that is (1000); 9 to 4, (0010); 13 to 6, (1100); 15 to 11, (0011) and the remaining patterns 7, (1010); 12, (1110); 14, (1011) and 16 (1111). These attribute patterns are identical to the patterns given in Table 3.

It can be easily verified that the reachability matrix given in earlier section (p. 13) is the same as the matrix which is obtained by examining the inclusion relationships among all combinations of the four column vectors of the attribute patterns in Table 3. This means that all possible knowledge states, obtainable from the four attributes with the structure represented by R can be used for diagnosing a student's errors. The five GRE items are good items as far as a researcher's interest is to measure and diagnose the nine states of knowledge and capabilities listed in Table 3.

Illustration With Real Examples

Example I: A Case of Discrete Attributes In Fraction Addition Problems

Birenbaum & Shaw (1985) used Guttman's facet analysis technique (Guttman, et.al. 1991) to identify eight task-content facets for solving fraction addition problems. There were six operation facets that described the numbers used in the problems and two facets dealing with the results. Then, a task specification chart was created based on a design which combined the content facets with the procedural steps. Figure 4 shows the task specification chart.

Insert Figure 4 about here

The task specification chart describes two strategies to solve the problems, methods A and B. Those examinees who use Method A convert a mixed number (a b/c) into a simple fraction, (ac+b)/c, similarly, the users of method B separate the whole number part from the fraction part and then add the two parts independently. In these cases, it is clear that when the numbers become larger in a fraction addition problem, then Method A obviously requires computational skills to get the correct answer. Method B, on the other hand, requires a deeper understanding of the number system.

Sets of attributes for the two methods are selected from the task specification chart in Figure 4 as follows:

Problem:	a b/c + d e/f	Method A	Method B
A_1	Convert (a b/c) to (ac+b)/c	used	Not used
A_2	convert (d e/f) to (df+e)/f	used	Not used
A_3	Divide fraction by a common factor	used	used
A_4	Find the common denominator of c & f	used	used
A_5	Make equivalent fractions	used	used
A_6	Add numerators	used	used
A_7	Divide numerator by denominator	used	used
A_8	Don't forget the whole number part	used	used
B_1	Separate a & d and b/c & e/f	Not used	used
B_2	Add the whole numbers including 0	Not used	used

The two methods share all of the attributes in common, except for B_1 and B_2, A_1 and A_2. The incidence matrices for the ten items in Birenbaum and Shaw (1985), for Methods A and B, are given in Table 4.

Insert Table 4 about here

A computer program written by Varadi and Tatsuoka (BUGLIB, 1990) produces a list of all the possible "can/cannot" combinations of attributes, otherwise known as the universal set of attribute response patterns.

For Method A, 13 attribute patterns are obtained. The attribute patterns and their corresponding ideal item patterns are given in Table 5 where the attributes are denoted by the numbers 1 through 8 for A_1 through A_8, and 9 and 10 for B_1 and B_2, respectively. For instance, the second state, 2, has the attribute pattern 11111110 and the ideal item pattern is represented by 111100010.

Insert Table 5 about here

It is interesting to note that there is no state including "cannot do an item that involves both of the attributes, A_1 and A_2, but can do items that involve either A_1 or A_2 alone" in the list given in Table 5. If one would like to diagnose such a compound state, then a new attribute should be added to the list.

Another interesting result is that A_5 cannot be separated from A_4 as long as we use only these ten items. In other words, the rows for A_4 and A_5 in the incidence matrix for Method A are identical. Needless to say, Shaw and Tatsuoka (1983) found many different errors that originated in attribute A_5, -- making equivalent fractions -- and they must be diagnosed for remediation (Bunderson & Ohlsen, 1983). In order to separate A_5 from A_4, we must add a new item which involves A_4 but not A_5, thereby making Row A_5 different from Row A_4.

Beyond asking the original "equivalent fraction" question, we now add an item to the existing item pool, which asks, "What is the common denominator of 2/5 and 1/7?" This is a way to test the skill for getting common denominators correctly and also distinguishes the separate skill required for making equivalent fractions. However, since the solutions to each of these questions a are so closely related and inter-dependent, it may not be possible to separate measure the examinees' skills in terms of each function.

If an examinee answers this item correctly but gets a wrong answer for items involving addition, such as 2/5 + 1/7, then it is more likely that the examinee has the skill for getting correct common denominators but not the skill for making equivalent fractions correctly.

Thirteen knowledge and capability states are identified from the incidence matrix for Method B, and they are also summarized in Table 5. Some ideal item response patterns can be found in the lists for both Methods A and B. This means that for some cases we cannot diagnose a student's underlying strategy for solving these ten items. Our attribute list cannot distinguish whether a student converts a mixed number (a b/c) to an improper fraction, or separates the whole number part from the fraction part. If we can see the student's scratch paper and can examine the numerators prior to addition, then we can find which method the student used. There are two solutions to this problem. One is to use a computer for testing so that crucial steps during problem solving activities can be coded. The second is to add new items so that these three attributes, A_1, A_2 and B_1 can be separated in the incidence matrix for Method B.

Example 2: The Case of Continuous and Hierarchically Related Attributes in The Adult Literacy Domain

Kirsch and Mosenthal (1990) have developed a cognitive model which underlies the performance of young adults on the so-called document literacy tasks. They identified three categories of variables which predict the difficulties of items with a multiple R of .94.

Three categories of variables are defined:

- "Document" variables (based on the structure and complexity of the document)
- "Task" variables (based on the structural relation between the document and the accompanying question or directive)
- "Process" variables (based on strategies used to relate information in the question or directive to information in the documents" (Kirsch and Mosenthal, 1990, p.5).

The "Document" variables comprise six specific variables including the number of organizing categories in the document, the number of embedded organizing categories in the document and the number of specifics. These three variables are considered in our incidence matrix as the attributes for "Document" variables.

The "Task" variables are determined on the basis of the structural relations between a question and the document that it refers to. The larger the number of units of information required to complete a task, the more difficult the task. Four attributes are picked up from this variable group.

The "Process" variables developed through Kirsch and Mosenthal's regression analysis showed that variables in the category of "Process" variables influenced the item difficulties to a large extent. One of the variables in this category is the degree of correspondence, which is defined as the degree to which the information given in the question or directive matches the corresponding information in the document.

The next variable represents the type of information which has to be developed to locate, identify, generate, or provide the requested information based on one or more nodes from a document hiererchy. Five hierarchically related attributes are determined from this variable group.

The last variables are Plausibility of Distractors, which measure the ability to identify the extent to which information in the document matches features in a question's given and requested information.

A total of 22 attributes are selected to characterize the 61 items. Since the attributes in each variable group are totally ordered, i.e., A_1 -> A_2 -> A_3 -> A_4 -> A_5, the number of possible combinations of "can/cannot" attributes is drastically reduced (Tatsuoka, 1991). One-hundred fifty-seven possible attribute response patterns were derived by the BUGLIB program and hence 157 ideal item response patterns are produced. As was explained in the earlier section, these 157 ideal item response patterns correspond to the 157 state distributions that are multivariate normal. These states are used for classifying an individual examinee's response pattern. A sample of ten states with their corresponding attribute response patterns are shown in

Table 6 as examples.

Insert Table 6 about here

As can be seen in Table 6, several subsets of attributes are totally ordered and the elements of the subset form a chain. Further 1500 subjects were classified into one of the 157 misconception states by a computer program entitled RULESPACE (Tatsuoka, Baillie, Sheehan, 1991). The number of subjects who were classified into one of these ten states are -- 157 subjects in State No.1, 46 in No. 4, 120 in No. 11, 81 in No. 12, 37 in No. 14, 68 in No. 50, 12 in No. 32, 27 in No. 102, 11 in No. 138 and 4 in No. 156.

While the interpretation of misconceptions for these results is described in detail elsewhere (Sheehan, Tatsuoka & Lewis, 1991), State No. 11 (into which the largest number of subjects were classified) will be described here.

"Cannot attributes A_{18} and A_{19}" relate directly from A_{18} to A_{19}. Therefore, as represented in Table 6, the statement can be made that, "a subject classified in this state cannot do A_{18}, and hence cannot, by default, do A_{19}." Thus, the prescription for these subjects' errors is likely to be that they make mistakes when items have the following specific feature:

....Distractors appear both within an organizing category and across organizing categories, because different organizing categories list the same specifics but with different attributes" (Kirsch and Mosenthal, 1990, p. 30).

Psychometric Theories Appropriate For A Constructed Response Format

An incidence matrix suggests various scoring formulas for the items.

First, the binary scores of right or wrong answers can be obtained from the condition that - if a subject can perform all the attributes involved in an item correctly, then the subject will get a score of one on that item; otherwise the subject will get a score of zero. With this scoring formula, the simple logistic models (Lord & Novick, 1968) for binary responses can be used for estimating the scaling variable θ.

Second, partial credit scores or graded response scores can be obtained from the incidence matrix if performance dependent on the attributes is observable and can be measured directly. This condition permits applicability of Masters' partial credit models (Masters, 1982) or Samejima's General Graded response models (Samejima, 1988) to data.

As far as error diagnoses are concerned, simple binary response models always work even when performances on the attributes <u>cannot</u> be measured directly and are <u>not</u> observable. However, computer scoring (Bennett, Rock, Braun, Frye, Spohrer, and Soloway, 1990), or scoring by human raters or teachers can assign graded scores to the items. For example, the number of correctly processed attributes for each item could be a graded score.

Muraki (1991) wrote a computer program for his modified version of Samejima's original graded response model (Samejima, 1969). Muraki's program can be used for Samejima's model itself also.

Third, a teacher may assign different weights to the attributes and give a student a score corresponding to the percentage of correct answers achieved, depending on how well the student performed on the attributes. Thus, the final score for the item becomes a continuous variable. Then Samejima's (1976, 1988) General Continuous IRT model can be used to estimate the ability parameter θ. If the response time for each item is available, then her Multidimensional Continuous model can be applied to such data sets.

Fourth, if a teacher is interested in particular combinations of attributes and assigns scores to nominal categories, say 1 = {can do A_1 and A_3}, 2 = {can do A_1 and A_2} and 3 = {can do A_2, A_3 and A_4},.. so on, then Bock's (1972) Polychotomous model can be utilized for getting θ.

Discussion

A wide variety of Item Response Theory models accommodating binary scores, graded, polychotomous, and continuous responses have been developed in the past two decades. These models are built upon a hypothetical ability variable θ. We are not against the use of global item scores and total scores -- e.g., the total score is a sufficient statistic for θ in the Rasch Model -- but it is necessary to investigate micro-level variables such as cognitive skills and knowledge and their structural relationships in order to develop a pool of "good" constructed- response items. The systematic item construction method enables us to measure unobservable underlying cognitive processes via observable item response patterns.

This study introduces an approach for organizing a couple of dozen such micro-level variables and for investigating their systematic interrelationships. The approach utilizes deterministic theories, graph theory and Boolean algebra. When most micro-level variables are not easy to measure directly, an inference must be made from the observable macro-level measures. An incidence matrix for characterizing the underlying relationships among micro-level variables is the first step toward achieving our goal. Then a Boolean algebra that is formulated on a set of sets of attributes, or a set of all possible item response patterns obtainable from the incidence matrix, enables us to establish relationships between two worlds: attribute space and item space (Tatsuoka, 1991).

A theory of item construction is introduced in this paper in conjunction with Tatsuoka's Boolean algebra work (1991). If a subset of attributes has a connected, directed relation and forms a chain, then the number of combinations of "can/cannot" attributes will be reduced dramatically. Thus, it will become easier for us to construct a pool of items by which a particular group of misconceptions of concern can be diagnosed with a minimum classification errors.

One of the advantages of rule space model (Tatsuoka, 1983,

1990) is that the model relates a scaled ability parameter θ to misconception states. For a given misconception state, which is error, one can always identify the particular types of errors which relate to ability level θ. If the centroid of the state is located in the upper part of the rule space, then one can conclude that this type of error is rare. If the centroid lies on the θ axis, then this error type is observed very frequently.

Although Rule space was developed in the context of binary IRT models, the concept and mathematics are general enough to be extended for use in more complicated IRT models. Further work to extend the rule space concept to accommodate complicated response models will be left for future research.

References

Bennett, R. E., Rock. D. A., Braun, H. I., Frye, D., Spohrer, J. C. & Soloway, E. (1990). The relationship of constrained free-response to multiple-choice and open-ended items. Applied Psychological Measurement, 14, 151-162.

Bennett, R. E., Ward, W.C., Rock, D. A., & LaHart, C. (1990). Toward a framework for constructed-response items (RR-90-7). Princeton, NJ: Educational Testing Service.

Birenbaum, M. & Shaw, D. J. (1985). Task Specification Chart: A key to better understanding of test results. Journal of Educational Measurement, 22, 219-230.

Birenbaum, M., & Tatsuoka, K. K. (1987). Open-ended versus multiple-choice response formats--it does make a difference for diagnostic purposes. Applied Psychological Measurement, 11, 329-341.

Bock, R. D. (1972). Estimating item parameters and latent ability when the responses are scored in two or more nominal categories. Psychometrika, 37, 29-51.

Brown, J. S., & VanLehn, K. (1980). Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 4, 370-426.

Bunderson, V. C., & Olsen, J. B. (1983). Mental errors in arithmetic: their diagnosis in precollege students. (Final Project Report, NSF SED 80-12500). WICAT, Provo, UT.

Chipman, S. F., Davis, C., & Shafto, M. G. (1986). Personnel and training research program: Cognitive science at ONR. Naval Research Review. 38, 3-21.

Dibello, L. V. & Baillie, R. J. (1991). Separating points in Rule space. (CERL Research Report). University of Illinois, Urbana, IL.

Easley, J. A. & Tatsuoka, M. M. (19680). Scientific thought, cases from classical physics. Allyn and Bacon, Boston.

Ginzburg, H. (1977). Children's arithmetic: The learning process. New York: Van Norstrand.

Glaser, R. (1985). The integration of instruction and testing. A paper presented at the ETS invitational Conference on the Redesign of Testing for the 21st Centry, New York, New York.

Guttman, R., Epstein, E. E., Amir, M., & Guttman, L. (1990). A structural theory of spacial abilities. Applied Psychological Measurement, 14, 217-236.

Hubert, L. J. (1974). Some applications of graph theory to clustering. Psychometrika, 39, 283-309.

Kim, S. H. (1990). Classification of item-response patterns into misconception groups. Unpublished doctoral dissertation. University of Illinois, Champaign.

Kirsch, I. S., & Mosenthal, P. B. (1990). Document literacy. Reading research quarterly, 25, 5-29.

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Masters, G. N. (1982). A Rasch model for partial credit scoring in objective tests. Psychometrika, 47, 149-174.

Muraki, E. (1991). Comparison of the graded and partial credit item response models. Unpublished manuscript. Princeton, NJ: Educational Testing Service, Princeton.

Samejima, F. (1969). Estimation of ability using a response pattern of graded scores. Psychometrika Monograph, 17.

Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. Psychometrika, 39, 111-121.

Samejima, F. (1988). Advancement of latent trait theory. ONR Final Report. University of Tennessee, Knoxville, Tenn.

Sato, T. (1990). An introduction to educational information technology. In Delwyn L. Harnisch & Michael L. Connell (Eds.), NEC Technical College, Kawasaki, Japan.

Sheehan, K., Tatsuoka, K. K., & C. Lewis (1991). Using the rule space model to diagnose document processing errors. A paper presented at the ONR conference, Workshop on Model-based Measurement, Educational Testing Service, Princeton, NJ.

Sleeman, D., Kelly, A. E., Martinak, R., Ward, R.,& Moore, J. (1989). Studies of diagnosis and remediation with high school algebra students. Cognitive Science, 13, 551-568.

Takeya, M. (1981). A study on item relational structure analysis of criterion referenced tests. Unpublished doctoral dissertation, Waseda University, Tokyo.

Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20, 345-354.

Tatsuoka, K. K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. Journal of Educational Statistics, 12, 55-73.

Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnoses. In N. Frederiksen, R. L. Glaser, A. M. Lesgold, & M. G. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition. Hillsdale, NJ: Erlbaum.

Tatsuoka, K. K. (1991). Boolean algebra applied to determination of the universal set of knowledge states. Technical Report-ONR-1, (RR-91-4). Princeton, NJ:Educational Testing Service.

Tatsuoka, K. K., Baillie, R. & Sheehan, K. (1991). RULESPACE: classifying a subject into one of the predetermined groups. Unpublished computer program.

Tatsuoka, K. K., Birenbaum, M., & Arnold, J. (1989). On the stability of students' rules of operation for solving arithmetic problems. Journal of Educational Measurement, 26, 351-361.

Tatsuoka, K. K., & Tatsuoka, M. M. (1987). Bug distribution and pattern classification. Psychometrika, 52, 193-206.

Tatsuoka, K. K., & Tatsuoka, M. M. (1991). On measures of Misconception stability. ONR-technical report. Princeton, NJ: Educational Testing Service, Princeton.

Tatsuoka, M. M. (1986). Graph theory and its applications in educational research: A review and integration. Review of Educational Research, 56, 291-329.

Tatsuoka, M. M., & Tatsuoka, K. K. (1989). Rule space. In S. Kotz and N. L. Johnson (Eds.), Encyclopedia of statistical sciences. New York: Wiley.

Varadi, F. & Tatsuoka, K. K. (1989). BUGLIB. Unpublished computer program. Trenton, New Jersey.

Warfield, J. N. (1973). On arranging elements of a binary in graphic form. IEEE transaction on systems, man and cybanetics, SMC-3, 121-132.

Warfield, J. N. (1973). Binary matrices in system modeling. IEEE transactions on systems, man and cybernetics, SMC-3, 441-449.

Wise, S. L. (1981). A modified order-analysis procedure for determining unidimensional items sets. Unpublished doctoral dissertation, University of Illinois, Champaign.

Acknowledgement

The author would like to gratefully acknowledge and thank several people for their help. Randy Bennett, Robert Mislevy, Kathy Sheehan, Maurice Tatsuoka, Bill Ward for valuable comments and suggestions, John Cordery for editorial help, Donna Lembeck for various help.

Table 1 A List of 16 Ideal Item Response Patterns Obtained from 16 Attribute Response Patterns by a Boolean Description Function

	Attribute response patterns	Ideal item response patterns
1	0000	1000000000000000
2	1000	1100000000000000
3	0100	1010000000000000
4	0010	1001000000000000
5	0001	1000100000000000
6	1100	1110010000000000
7	1010	1101001000000000
8	1001	1100100100000000
9	0110	1011000010000000
10	0101	1010100001000000
11	0011	1001100000100000
12	1110	1111011010010000
13	1101	1110110101001000
14	1011	1101101100100100
15	0111	1011100011100010
16	1111	1111111111111111

Table 2 A List of Attribute Response Patterns and Ideal Item Response Patterns Affected by Direct Relations of Attributes

Original Patterns	Attribute Patterns	Ideal Item Patterns
1,3,4,5,9,10,11,15	0000	1000000000000000
2, 7	1000	1100000000000000
8,14	1001	1100100100000000
13	1101	1110110101001000
6	1100	1110010000000000
12	1110	1111011010010000
16	1111	1111111111111111

Table 3 A List of Nine Knowledge and Capability States and Nine Ideal Item Patterns of GRE-math items

Attribute Patterns		Ideal Item Patterns	Description of States
1	1 1 1 1	1 1 1 1 1	Can do everything
2	1 1 1 0	0 1 1 0 1	Can do A_1*, A_2, A_3 Cannot do A_4
3	1 1 0 0	0 1 1 0 0	Can do A_1, A_2 Cannot do A_3, A_4
4	1 0 1 1	0 0 1 1 1	Can do A_1, A_3, A_4 Cannot do A_2
5	1 0 1 0	0 0 1 0 1	Can do A_1, A_3 Cannot do A_2, A_4
6	1 0 0 0	0 0 1 0 0	Can do A_1 Cannot do A_2, A_3, A_4
7	0 0 1 1	0 0 0 1 1	Can do A_3, A_4 Cannot do A_1, A_2
8	0 0 1 0	0 0 0 0 1	Can do A_3 Cannot do A_1, A_2, A_4
9	0 0 0 0	0 0 0 0 0	Cannot do anything

* A_1 : Goal is to find the net filling rate

A_2 : Compute the rate

A_3 : Goal is to find the time to fill the tank

A_4 : Compute the time.

Table 4 Ten Items with Their Attribute Characteristics by Method A and Method B

Method A

1	2 8/6 + 3 10/6	A_1, A_2, A_3, A_6, A_7
2	3/5 + 1/5	A_6
3	3 10/4 + 4 6/4	A_1, A_2, A_3, A_6, A_7
4	7/4 + 5/4	A_6, A_7
5	3/4 + 1/2	A_4, A_5, A_6, A_7, A_8
6	2/5 + 12/8	A_3, A_4, A_5, A_6, A_7, A_8
7	1/2 + 1 10/7	A_2, A_4, A_5, A_6, A_7, A_8
8	1/3 + 1/2	A_4, A_5, A_6
9	3 1/6 + 2 3/4	A_1, A_2, A_4, A_5, A_6, A_7, A_8
10	5/6 + 1/3	A_4, A_5, A_6, A_7, A_8

Method B

1	2 8/6 + 3 10/6	B_1, A_3, A_4, A_5, A_6, A_7, B_2
2	3/5 + 1/5	same as by Method A
3	3 10/4 + 4 6/4	B_1, A_3, A_6, A_7, A_8, B_2
4	7/4 + 5/4	same as by Method A
5	3/4 + 1/2	same as by Method A
6	2/5 + 12/8	same as by Method A
7	1/2 + 1 10/7	B_1 A_4, A_5, A_6, A_7, A_8, B_2
8	1/3 + 1/2	same as by Method A
9	3 1/6 + 2 3/4	B_1, A_4, A_5, A_6, B_2
10	5/6 + 1/3	same as by Method A

Table 5 A list of all the possible sets of attribute patterns derived from the incidence matrices given in Table 4

Method A

States	Cannot	Can	Ideal Item Response Pattern
1	none	1,2,3,4,5,6,7,8	1111111111
2	8	1,2,3,4,5,6,7	1111000100
3	4,5,8	1,2,3,6,7	1111000000
4	1	2,3 4,5,6,7,8	0101111101
5	2,1	3,4,5,6,7,8	0101110101
6	3	1,2,4,5,6,7,8	0101101111
7	3,1	2,4,5,6,7,8	0101101101
8	3,2,1	4,5,6,7,8	0101100101
9	1,2,3,8	4,5,6,7	0101000100
10	1,2,3,4,5,8	6,7	0101000000
11	7,1,2,3,8	4,5,6	0100000100
12	1,2,3,8,7,4,5	6	0100000000
13	1,2,3,4,5,6,7,8	none	0000000000

Method B

States	Cannot	Can	
1	none	3,4,5,6,7,8,9,10	1111111111
2	8	3,4,5,6,7,9,10	1101000110
3	4,5	3,6,7,8,9,10	0111000000
4	9,10	3,4,5,6,7,8	0101110101
5	3	4,5,6,7,8,9,10	0101101111
6	3,9,10	4,5,6,7,8	0101100101
7	3,8	4,5,6,7,9,10	0101000110
8	3,8,9,10	4,5,6,7	0101000100
9	3,4,5,8,9,10	6,7	0101000000
10	7,3 8	4,5,6,9,10	0100000110
11	3,7,8,9,10	4,5,6	0100000100
12	3,4,5,7,8,9,10	6	0100000000
13	3,4,5,6,7,8,9,10	none	0000000000

Table 6 The Ten States Selected from One-hundred Fifty-seven Possible States Yielded by Boolean Operation (via BUGLIB program)

	States	Attribute Pattern 1111111111222 1234567890123456789012	Directed Direct Relation Among Attributes
1	No. 1	1111111111111111111111	None
2	No. 4	1111111111111111110111	None
3	No. 11	1111111111111111100111	$A_{18} \rightarrow A_{19}$
4	No. 12	1111011111111111100111	$A_{18} \rightarrow A_{19}$
5	No. 14	1111011110111111100111	$A_{18} \rightarrow A_{19}$
6	No. 30	1111011100111111100111	$A_9 \rightarrow A_{10}$, $A_{18} \rightarrow A_{19}$
7	No. 32	1100011100111111100110	$A_3 \rightarrow A_4 \rightarrow A_5$, $A_9 \rightarrow A_{10}$
8	No. 102	1000011111111111111111	$A_2 \rightarrow A_3 \rightarrow A_4 \rightarrow A_5$
9	No. 138	1000011111111011110111	$A_2 \rightarrow A_3 \rightarrow A_4 \rightarrow A_5$
10	No. 156	1000010000001110000100	$A_2 \rightarrow A_3 \rightarrow A_4 \rightarrow A_5$ $A_7 \rightarrow A_8 \rightarrow A_9 \rightarrow A_{10}$ $A_{11} \rightarrow A_{12} \rightarrow A_{13}$ $A_{16} \rightarrow A_{17} \rightarrow A_{18} \rightarrow A_{19}$ $A_{21} \rightarrow A_{22}$

A systematic analysis of

task

skill

job

content

identifying prime components, abstracting attributes and naming them $A_1, \ldots\ldots, A_k$.

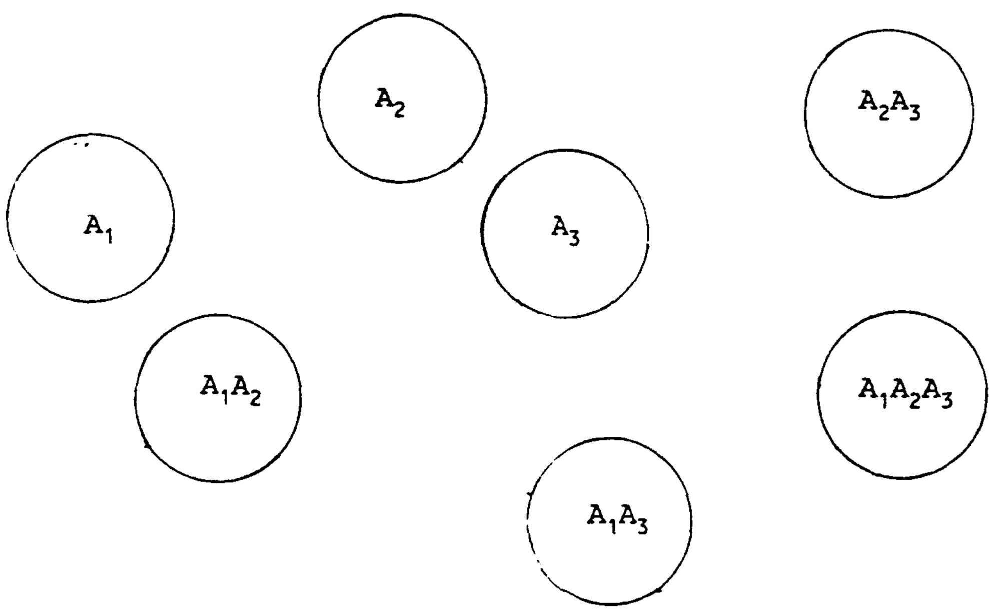

Figure 1 Examples of Attributes

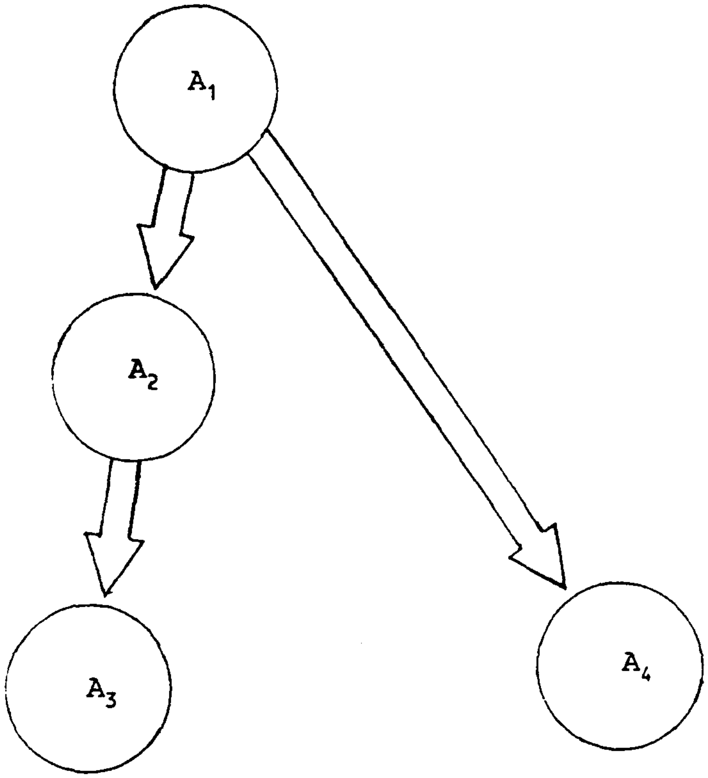

Figure 2 An Example of Partially Ordered Attributes

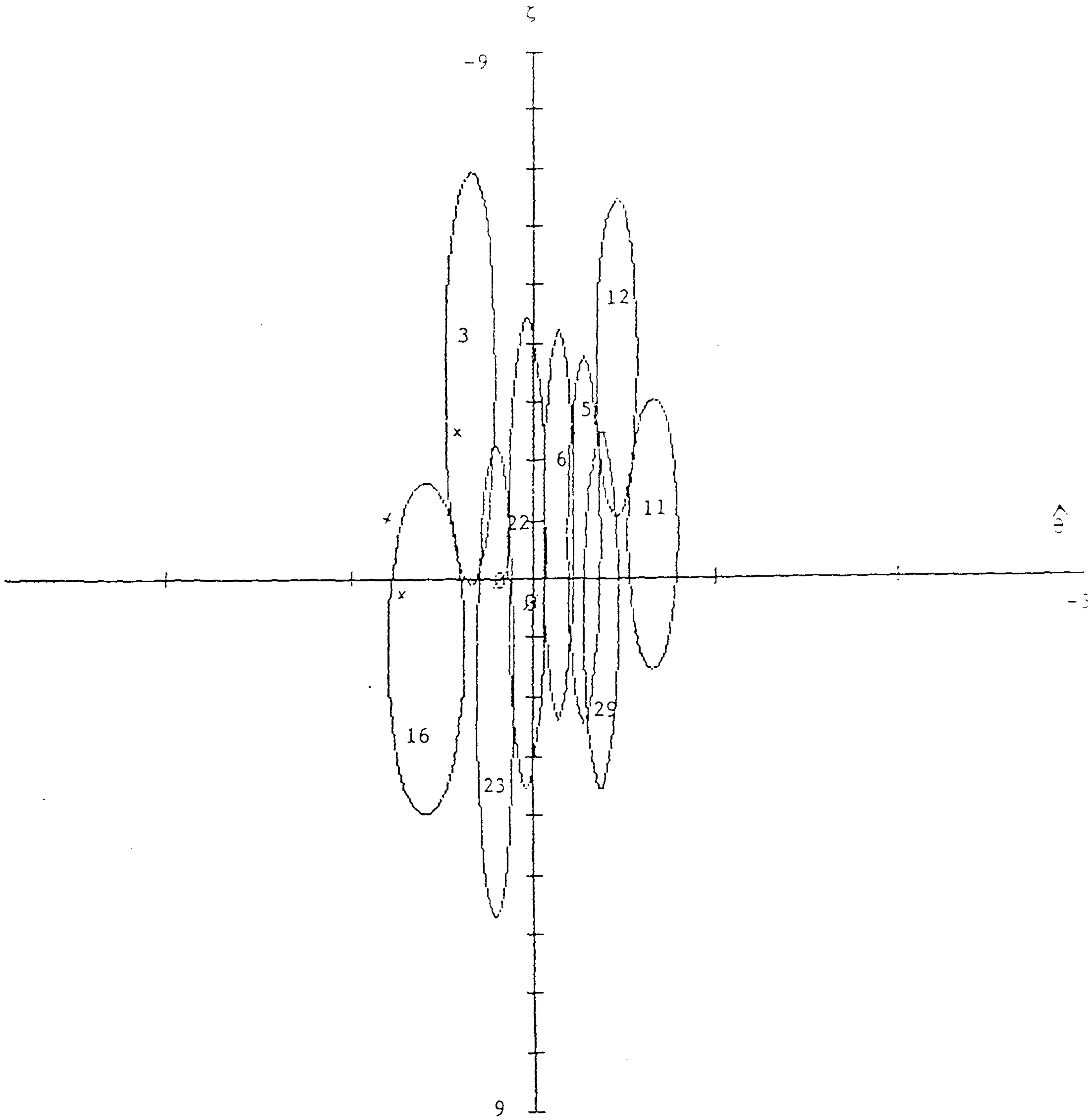

Figure 3   The Rule Space Configuration.

The Numbers in Nine ellipses indicate error States (e.g., No. 5 State is "one cannot do the operation of borrowing in fraction subtraction problems.") and x marks represent students' points (θ, ζ).

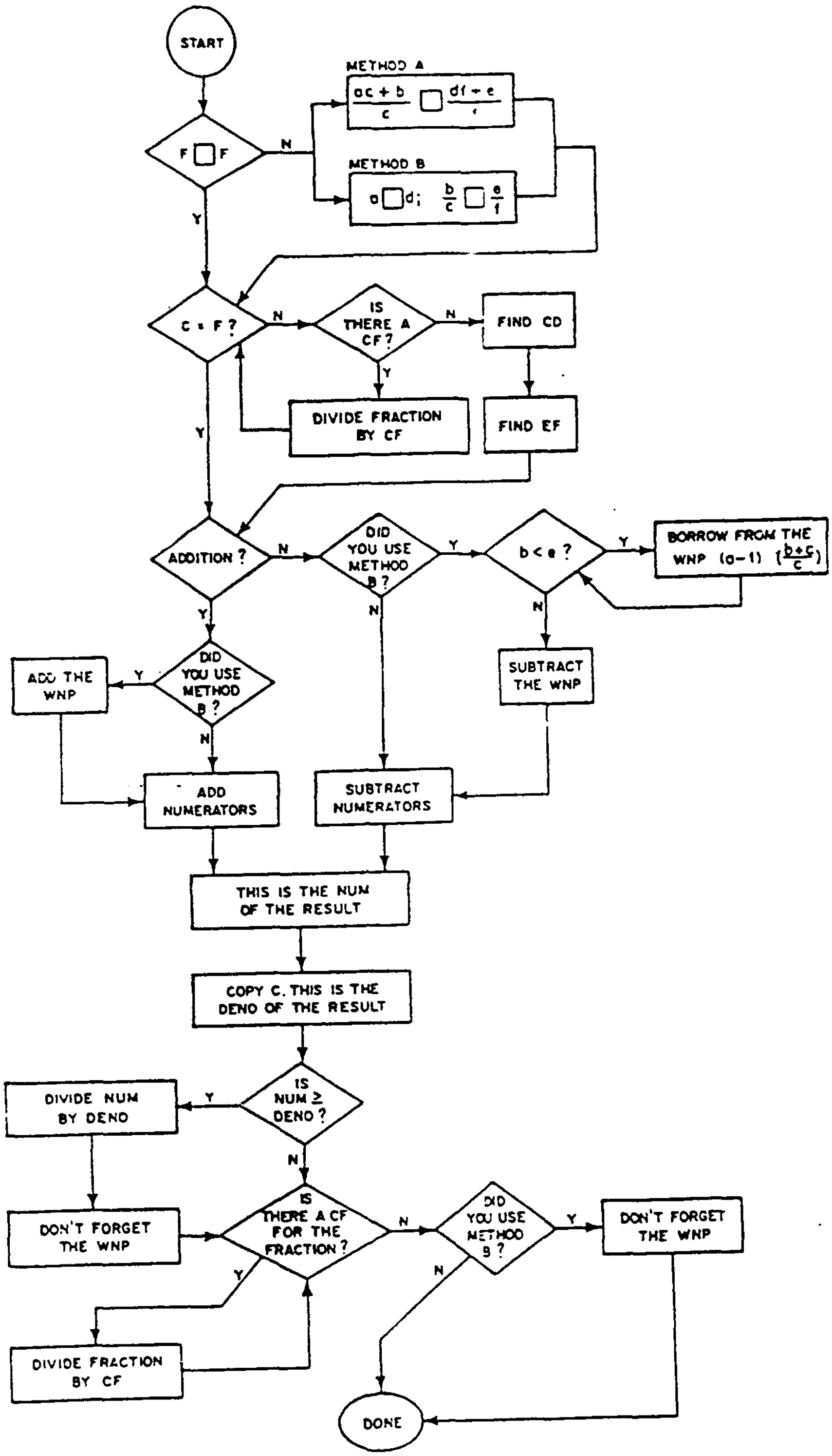

Figure 4 Task Specification Chart for Fraction Addition and Subtraction Problems.

Symbol used to denote the general fraction form used in this figure is: a(b/c) + d(e/f); F is fraction; CD is common denominator; CF is common factor; WNP is whole number part; NUM is numerator; DENO is denominator; EF is equivalent fraction.

Distribution List

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. James Algina
1403 Norman Hall
University of Florida
Gainesville, FL 32605

Dr. Nancy Allen
Educational Testing Service
Princeton, NJ 08541

Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Isaac Bejar
Law School Admissions
Services
P.O. Box 40
Newtown, PA 18940-0040

Dr. Anne Beland
Educational Testing Service
Princeton, NJ 08541

Dr. Ira Bernstein
Department of Psychology
University of Texas
P.O. Box 19528
Arlington, TX 76019-0528

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Ramat Aviv 69978
ISRAEL

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Gwyneth Boodoo
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Breaux
Code 252
Naval Training Systems Center
Orlando, FL 32826-3224

Dr. Robert Brennan
American College Testing
Programs
P. O. Box 168
Iowa City, IA 52243

Dr. David V. Budescu
Department of Psychology
University of Haifa
Mount Carmel, Haifa 31999
ISREAL

Dr. Gregory Candell
CTB/McGraw-Hill
2500 Garden Road
Monterey, CA 93940

Dr. John B. Carroll
409 Elliott Rd., North
Chapel Hill, NC 27514

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute, H1-B52
P.O. Box 704
Yorktown Heights, NY 10598

Dr. Robert M. Carroll
Chief of Naval Operations
OP-01B2
Washington, DC 20350

Dr. W. Chambers
Technology Manager, Code 2B
Naval Training Systems Center
12350 Research Parkway
Orlando, FL 32826-3224

Mr. Hua Hua Chang
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Raymond E. Christal
UES LAMP Science Advisor
AFHRL/MOEL
Brooks AFB, TX 78235

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
Los Angeles, CA 90089-1061

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
4401 Ford Avenue
Alexandria, VA 22302-0268

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans F. Crombag
Faculty of Law
University of Limburg
P.O. Box 616
Maastricht
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

Dr. C. M. Dayton
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Ralph J. DeAyala
Measurement, Statistics,
and Evaluation
Benjamin Bldg., Rm. 4112
University of Maryland
College Park, MD 20742

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Neil Dorans
Educational Testing Service
Princeton, NJ 08541

Dr. Fritz Drasgow
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
(2 Copies)

Dr. Stephen Dunbar
224B Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. George Englehard, Jr.
Division of Educational Studies
Emory University
210 Fishburne Bldg.
Atlanta, GA 30322

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Leonard Kroeker
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Jerry Lehnus
Defense Manpower Data Center
Suite 400
1600 Wilson Blvd
Rosslyn, VA 22209

Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

Dr. Richard Lesh
Educational Testing Service
Princeton, NJ 08541

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Educational Testing Service
Princeton, NJ 08541-0001

Ms. Hsin-hung Li
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert L. Linn
Campus Box 249
University of Colorado
Boulder, CO 80309-0249

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. George B. Macready
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen J. Martin
Office of Chief of Naval
Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

Dr. Shin-ichi Mayekawa
The National Center for University
Entrance Examinations
2-19-23 KOMABA, MEGURO-KU
Tokyo 153
JAPAN

Dr. James R. McBride
HumRRO
6430 Elmhurst Drive
San Diego, CA 92120

Dr. Clarence C. McCormick
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Dr. Michael McNeese
DET-1, AL/CFHI
BLDG 248
Wright-Patterson AFB, OH 45432

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152-6800

Ms. Kathleen Moreno
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Librarian
Naval Center for Applied Research
in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. James B. Olsen
WICAT Systems
1875 South State Street
Orem, UT 84058

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Dr. Judith Orasanu
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MOA
Brooks AFB, TX 78235

Mr. Steve Reiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. W. A. Rizzo
Head, Human Factors Division
Naval Training Systems Center
Code 26
12350 Research Parkway
Orlando, FL 32826-3224

Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Mr. Louis Roussos
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37916-0900

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

Mr. Kenneth Sarno
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Janice Scheuneman
Educational Testing Service
Princeton, NJ 08541

Lowell Schoer
Psychological & Quantitative Foundations
College of Education
University of Iowa
Iowa City, IA 52242

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

Dr. Dan Segall
Navy Personnel R&D Center
San Diego, CA 92152

Mr. Robert Semmes
N218 Elliott Hall
Department of Psychology
University of Minnesota
Minneapolis, MN 55455

Dr. Robin Shealy
Illinois State Water Survey
Room 149
2204 Griffith Dr.
Champaign, IL 61820

Ms. Kathleen Sheehan
Educational Testing Service
Princeton, NJ 08541

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujisawa 251
JAPAN

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. Richard C. Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Judy Spray
ACT
P.O. Box 168
Iowa City, IA 52243

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. William Stout
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
Code-62
San Diego, CA 92152-6800

Dr. John Tangney
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Maurice Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Mr. Sherman Tsien
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Robert Tsutakawa
University of Missouri
Department of Statistics
222 Math. Sciences Bldg.
Columbia, MO 65211

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 440
St. Paul, MN 55114

Dr. Frank L. Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman
Box 146
Carmel, CA 93921

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 51
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Rand R. Wilcox
University of Southern California
Department of Psychology
Los Angeles, CA 90089-1061

German Military Representative
ATTN: Wolfgang Wildgrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

Dr. David Wiley
School of Education
Northwestern University
Evanston, IL 60201

Dr. Charles Wilkins
Navy Personnel R&D Center
Code 13
San Diego, CA 92152

Dr. Bruce Williams
Department of Educational Psychology
University of Illinois
Urbana, IL 61801

Dr. Mark Wilson
School of Education
University of California
Berkeley, CA 94720

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, NY 10021