MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. MEMO - 488
LOGO MEMO - 50

August 1978

UNDERSTANDING UNDERSTANDING MATHEMATICS

by

EDWINA RISSLAND MICHENER

Abstract

In this paper we look at some of the ingredients and processes involved in the understanding of mathematics. We analyze elements of mathematical knowledge, organize them in a coherent way and take note of certain classes of items that share noteworthy roles in understanding. We thus build a conceptual framework in which to talk about mathematical knowledge. We then use this representation to describe the acquisition of understanding. We also report on classroom experience with these ideas.

The work reported in this paper was supported in part by the National Science Foundation (under grant number 77-19083SEI) and conducted at the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts. The views and conclusions contained in this paper are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the National Science Foundation or the United States Government.

Table of Contents

1. Introduction

When a mathematician says he understands a mathematical theory, he possesses much more knowledge than that which concerns the deductive aspects of theorems and proofs. He knows about examples and heuristics and how they are related. He has a sense of what to use and when to use it, and what is worth remembering. He has an intuitive feeling for the subject, how it hangs together, and how it relates to other theories. He knows how not to be swamped by details, but also to reference them when he needs them.

This paper is concerned with this important extra-logical knowledge that is often outside of traditional discussions in mathematics. The goal is to develop a conceptual framework in which to talk about mathematical knowledge and to understand the understanding of mathematics, in order to improve how to learn, teach, and do mathematics.

Polya once remarked that, "A well-stocked and well-organized body of knowledge is an asset to the problem solver. Good organization which renders the knowledge readily available may be even more important than the extent of the knowledge." [18, p.85]. The same is true for aspects of mathematics other than problem solving. Thus, our first task is to seek answers to the questions: "What are the ingredients of mathematical knowledge, their types and their functions?", "How can this knowledge be organized and represented?"

2. An Epistemology of Mathematical Knowledge

This section presents a conceptual framework for mathematical knowledge that is based on the role of various kinds of knowledge in the understanding of mathematics in general and of mathematical theories in particular. This epistemology is based on material found in textbooks, lecture notes, discussions, and protocols of neophyte and expert mathematicians; it reflects the experience of teachers, students and practitioners of mathematics. Where there were alternatives for classifying and representing knowledge, the one choosen was that that best fit the author's experiences in learning and in helping students to learn mathematics. The reader is cautioned that this epistemology is not complete and exhaustive, that is, it does not represent all aspects of mathematical knowledge.

When one analyzes the knowledge that mathematicians -- students and professionals -- use when they do and explain mathematics, it becomes clear that there are several kinds of mathematical knowledge: (1) clusters of strongly bound pieces of information, such as the statement of a theorem, its name, its proof[s], an evaluation of its importance, which can be taken together to comprise a single *item*, such as a theorem; and (2) relations between the items, such as the logical connections between theorems.

2.1 Examples, Results and Concepts

We can distinguish three major categories of items: *results*, which contain the traditional logical-deductive elements of mathematics, i.e., theorems; *examples*, which contain illustrative material; and *concepts*, which contain mathematical definitions and heuristic notions and advice.

Results can be organized by the relation of *logical support* in which A --> B means that result A is used to prove result B. For instance in the theory of unique factorization [8], before one can prove that every integer can be uniquely factored, one must first prove supporting results on the Euclidean algorithm and the greatest common divisor. Results together with the relations of logical support comprise *Results-space*.

The collection of examples also has a natural relation. Examples can be organized by the relation of *constructional derivation* in which A --> B means that example A is used to construct example B. For instance, the Cantor set is constructed from the unit interval by the process of "deleting middle thirds" [7, 22], and thus, *unit interval --> Cantor set*:

0____________________1 *the unit interval*

define sequence of sets by deleting middle thirds

0______(   )______1

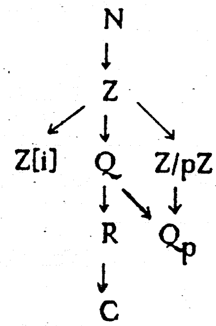
0__( )__(   )__( )__1

0_()_( )_()_(   )_()_( )_()_1

limiting set is the Cantor set

Cantor functions and higher dimensional Cantor sets can also be constructed from the unit interval and the Cantor set [4, 7, 22].

The following cluster of examples is familiar to everyone's arithmetic experience. One starts with the natural numbers N, and then builds the integers Z by introducing non-positive integers, say, by doing subtractions. Quotients then lead to the rational numbers Q, from which one can build the real numbers R, by filling in the "gaps". If one goes on to study elementary number theory, one then builds more examples, such as the Gaussian integers Z[i], the field of integers modulo a prime Z/pZ, and the p-adic numbers Q_p. These examples can be organized according to their constructional relations:



The p-adic numbers have arrows coming from both Q and Z/pZ since either can be used to construct Q_p. (Q_p can be constructed from Q by completion of Cauchy sequences with respect to a metric just like the construction of the reals R from Q, or from Z/pZ by an algebraic construction involving "inverse limits" [2]). The point is that there are two constructional routes leading to Q_p, and thus a directed graph, and not simply a tree, is needed to show the relations.

Examples have some general properties worth noting: (1) pictures are an integral part of many examples; (2) constructions are like procedures; (3) the pictures need not be static; in fact those shown for the Cantor set are merely a few frames from a sequence.

The category of concepts includes both formal and informal ideas, that is, definitions and heuristics. Definitions could alternatively be included in Results-space since they are the logical atoms from which the proofs are deduced. This approach was not adopted; rather, they are included in Concepts-space to highlight the interdependence of ideas and keep track of conceptual dependencies as one's network of ideas grows. Informal ideas, e.g., "mega-principles" and "counter-principles" (see Section 2.3), often evolve from more formal ones, e.g., definitions, under the forces of such "genetic" processes [5, 14] as paraphrase [12], analogy, generalization, specialization [17], and "monster-barring" [10].

Concepts can be organized by the pedagogical judgement that concept A should be known about before concept B, which we shall call the relation of *pedagogical ordering*. Sometimes it simply reflects the fact that concept A enters into the definition of concept B; at other times, it reflects expository tastes. For instance in studying arithmetic properties of the integers, one needs to know about division before being able to talk about primes; once one knows about primes, one can go on to discuss prime factorizations.

A concept can be expressed either as a declarative statement -- the familiar formulation of most mathematical definitions -- or as a procedure or the result of a procedure. Some concepts are most naturally expressed in declarative form, and others, such as the Gram-Schmidt process, Gaussian elimination, Newton's method, are most naturally expressed as procedures. Some concepts can be expressed in either way, such as "eigenvalue" which can be defined either as the λ of $Av=\lambda v$, or as a root of the characteristic polynomial $\det(A - \lambda I)$

= 0. The concept of "square root of a number" has aspects involving both declarative and procedural knowledge: it can be thought of either as an "x" whose square is that number, or as the outcome of an algorithm like that taught in high school.

Thus, mathematical knowledge can be structured by three major types of item/relation pairs -- *examples/constructional derivation, results/logical support, and concepts/pedagogical ordering* -- which establish three representation spaces for a mathematical theory: *Examples-space, Results-space* and *Concepts-space.* They are best shown as directed graphs where the direction matches the predecessor-successor ordering inherent in the relations.

2.2 Dual Relations [1]

When considering a theory item, one can decide whether to classify it as a result, example or concept and then fit it into the appropriate representation space by determining its predecessors and successors. One can also consider items outside of the representation space with which it is associated. *Dual relations* concern these inter-space relations. They are introduced to capture both the way in which one's attention moves easily between the three types of items, and the naturalness with which we associate items that often seem to be distantly related in the senses defined by the in-space relations.

Specifically, *dual items* are defined as follows:

> *The dual items of an example are: the ingredient concepts and results needed to discuss or construct it, and the concepts and results motivated by it.*
>
> *The dual items of a result consist of: the examples motivating it, the concepts needed to state and prove it, and the concepts and examples that are derived from it.*
>
> *The dual items of a concept are: the examples motivating it, the results laying the groundwork for it; and the examples illustrating it and the results proving things about it.*

Thus the *dual* of an item contains sets of the other two kinds of items:

dual(an example) = {*results*}, {*concepts*}
dual(a result) = {*examples*}, {*concepts*}
dual(a concept) = {*examples*}, {*results*}

[1] The use of the word dual here has no technical relation to its use in the theory of vector spaces, although there is a metaphorical connection.

The subset of examples in the dual set of an item is called the *examples-dual*, the subset of results, the *results-dual*, and the subset of concepts, the *concepts-dual*.

The dual of an item can alternatively be sub-divided into subsets containing items that precede the item in the understanding or development of a theory, the *pre-dual*, those that come after the item, the *post-dual*, and those that have neither a strong "pre" nor "post" flavor. To use Polya's words, the pre-dual items are "suggestive", and the post-dual, "supportive" of an item [17, pp. 4-5]. For instance, concepts needed to prove a theorem would be included in the pre-concepts-dual of the result.

Two items are said to be *related via the dual idea* if they share common dual items. The mathematical world is full of dual relations: the examples of the real numbers and the p-adic numbers are related via the concept of completion; the concepts of measure and length via the example of an interval (a,b); Pythagoras's Theorem and the Law of Cosines, via an example of a right triangle; concepts of continuity and differentiability, via the example of the absolute value function; concepts of countability and measure zero are related via the Cantor set; concepts of fixed points and the power of an operation, via the example of $\cos^n x$.

One can define various equivalence relations that are based upon the dual idea. For instance, two items are *dual equivalent* if their duals are equal. Two dual equivalent results would share identical sets of concepts and examples, for instance. Dual equivalent items are very similar and in many senses are the "same" and should be identified.

Relation via the dual idea is extremely useful because it describes how we associate knowledge that is not necessarily closely related in the sense of the three relations operating within the three representation spaces. It also ties the spaces back together.

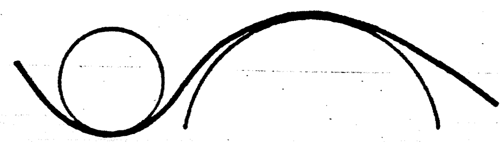
2.3 Epistemological Classes

Not all examples, results and concepts are equally important or serve the same function in one's understanding. We group those that play noteworthy roles in understanding into *epistemological classes*. These classes are not necessarily disjoint since an item may play more than one role.

For instance, when learning a theory for the first time, one can grasp certain perspicuous examples immediately and easily. These *start-up* examples help one get started in a new subject by motivating basic definitions and results, and setting up useful intuitions.

The "circles and lines" example is such a start-up example from differential geometry. The following is a paraphrase of Spivak's use of it [24] to introduce the theory of curvature of plane curves:

We begin by considering circles and lines. We can agree that circles curve and that lines don't. Furthermore, small circles "curve more" than large circles. (This is consistent with our observations about lines, which are a limiting case.) Thus we note that curvature is inversely related to the radius. So for a circle we define the curvature to be the reciprocal of its radius. Now what about more general plane curves? Well, we *lift* our circle-line definition to the general case by fitting circles onto the curves:



This simple example suggests how one can approach the study of curvature; when formalized, this example becomes the osculating circle definition of curvature. It provides a strong pictorial representation for curvature (circles) and a handle (the osculating circle) for calculating it.

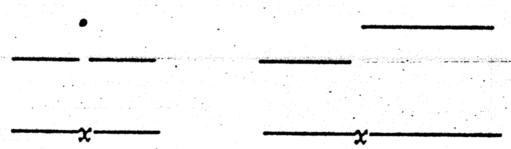
This example exhibits many properties that a start-up example should have: (1) it motivates fundamental concepts; (2) it can be understood by itself; (3) it is *projective*, i.e., its specific situation can be lifted to the general case; and (4) it provides a simple and suggestive picture.

Reference examples are another important class of examples. They are examples that one refers to over and over again. They are basic, widely applicable and provide a common point of contact through which many results and concepts are linked together.

Reference examples are used as standard cases to check out one's understanding. For instance, no matter where one is in the study of real analysis, one invariably refers to R^2 to see how things really work. In elementary number theory as well as algebraic number theory, one always looks at the integers Z. In linear algebra, a very useful reference example is what Michener [11] calls the *Basic 16*; it is the collection of the sixteen 2X2 matrices whose entries are 0's and 1's. This collection contains examples illustrating many of the "good" as well as the "bad" things that occur in eigenanalysis; it is also a rich source for counter-examples. These matrices seem to be ubiquitous throughout mathematics.

Model examples are paradigmatic, generic examples. They suggest and summarize expectations and default assumptions about results and concepts. They are indicative of the general case.

For instance, in the study of real-valued functions, the following diagrams indicate the general kind of behavior a function has at a point where it has a simple discontinuity. The diagram on the left represents a function with an "aberration" discontinuity at x, i.e., the right and left hand limits exist and are the same, but the function has the "wrong" value at x, and the diagram on the right represents a "jump" discontinuity, i.e., right and left limits exist but are not the same:



Observe that the specific measurements in these pictures are unimportant; what counts is that they capture the essence of the situation.

Because of their generic nature, model examples are often closely related to *without loss of generality* arguments. For instance, the model examples for conic sections are usually pictured as having their major axes aligned with the x- and y-axes (see any calculus book, e.g., [26]); these diagrams are completely general because one can always use the coordinate transformations of translation and rotation to change variables in order that the axes are so aligned.

Model examples are flexible and manipulatable structures which usually must be fine-tuned to meet the specifics of a problem. For instance, to capture the fact that a function has a "big" jump discontinuity, the lines in the above example could be made very far apart.

Counter-examples are familiar to everyone as examples that show a statement is not true. They sharpen distinctions between concepts.

Some counter-example are referenced frequently. For instance, the Cantor set is used repeatedly in the study of measure and integration as a counter-example in connection with items whose concepts-dual includes the concept of "measure zero". A specific use is with the result "countable sets have measure zero" whose converse is refuted by the Cantor set which is an example of an uncountable set that has measure zero [7, 21, 22]. The factorization of 2 as (1-i)(1+i) is often used as a counter-example to show that not all rational primes (i.e., numbers which are primes in Z) are prime in the setting of the Gaussian integers Z[i].

Other counter-examples are used once to establish a point and then are abandoned. Such a *hapax legomenon* [3] has a limited use in the theory and memory of it is often very short-lived, perhaps because it has so few connections to the rest of one's knowledge.

In summary, major epistemological classes of Examples-space are: *start-up examples, reference examples, model examples*, and *counter-examples.*

Concepts-space has two major epistemological classes in addition to the obviously important class of definitions. These other classes contain the heuristic advice that we give to ourselves and to others, while working in a theory [2]: mega-principles and counter-principles.

Mega-principles (*MP*'s) are kernels of wisdom in the form of powerful suggestions or generally valid statements. For instance, the MP: *Look at extreme points* is a very powerful heuristic in calculus and analysis. *Symmetric matrices are nice* is a mega-principle from linear algebra; it is a synopsis of many results that show symmetric matrices are well-behaved, e.g., diagonalizable and numerically stable [13, 25]. *Try the 2X2 case* is powerful advice in the study of matrices. Another useful suggestion in this and other domains is the MP to *Try special cases involving only 0's and 1's.*

Royden's analysis book presents "Littlewood's Three Principles" which are striking examples of mega-principles; Royden quotes from Littlewood [21, Chapter 3, Section 6]:

> "There are three principles, roughly expressible in the following terms: Every (measurable) set is nearly a finite union of intervals; every (measurable) function is nearly continuous; every convergent sequence of (measurable) functions is nearly uniformly convergent. Most of the results are fairly intuitive applications of these ideas... If one of the principles would be the obvious means to settle the problem if it were 'quite' true, it is natural to ask if the 'nearly' is near enough, and for a problem that is actually solvable it generally is."

In summary some MP's provide imperatives or advice while others give an idea of what to expect. Mega-principles express broad "flavors" of a theory that are often remembered long after the details have been forgotten. Like model examples, they provide broad, suggestive, initial descriptions and expectations.

Counter-principles (*CP*'s) alert one to possible sources of blunders or troubles. For instance, everyone knows about the CP: *Watch out for division by 0.* In linear algebra and numerical analysis, the counter-principle *Multiple roots are troublesome* warns of potential trouble when multiple roots occur, e.g., in diagonalizing or numerical computations. The CP from calculus -- *when changing the variable of integration, don't forget to calculate the new differential:* $dv=v'(x)dx$ -- is a word of warning familiar to all calculus students.

[2] Polya [16] and Schoenfeld [23] deal with more general, domain-independent strategies, whereas our concern is with heuristics that are relevant to a particular domain, although some might indeed be useful in a larger context.

CP's are distillations of many results, counter-examples, and failed attempts. Like counter-examples they add focus and limits to one's intuitions. They are often related to MP's as cavaets to warn of misapplications of the MP. For instance, related to the MP(n=2) is the CP that suggests being careful about jumping to (inductive) conclusions without checking out the case of n=3.

Results-space also has several epistemological classes.

Basic results establish elementary but important properties of concepts and examples. For example, the result: *λ is an eigenvalue of the matrix A (i.e., $Av=\lambda v$) iff $det(A-\lambda I)=0$* is a result basic to the study of eigenvalues. It relates the procedural formulation of the eigenvalue concept (solving the characteristic equation) with the declarative (existential) formulation. Other basic results link concepts with examples, such as *The outer measure of an interval is its length* which links the concepts of measure and length via the reference example of an interval as well as relating the concept of measure to the interval example.

Key results establish fundamental facts of a theory which are used repeatedly once they have been proved. For instance, the "Side-Angle-Side" theorem is a key result from plane geometry.

Culminating results are the goal results towards which a theory drives. The test of a culminating result is to ask, "If this result is omitted has the main point of the theory been missed?" If the answer is yes, the result is a culminating result. For example, the Fundamental Theorem of Calculus is a culminating result from calculus. The Jordan Normal Form Theorem, the Cauchy Integral Formula, the Riesz Representation Theorem, and many other "name" results are culminating results of their theories. Many culminating results are equivalency or classification results that connect alternative descriptions, definitions and approaches, such as the theorem showing that all real vector spaces of a given dimension are isomorphic, or Wedderburn's Theorem which gives a large number of different formulations of projective modules [9].

Less important than basic, key and culminating results are *transitional* and *technical* results which provide logical stepping-stones and work out technical details for a theory.

There are many analogies between the epistemological classes: model examples, mega-principles, and culminating results are all important items within their categories which are usually remembered for a long time; counter-examples and counter-principles serve a limiting function; basic results and start-up examples provide easy starting points in a theory; reference examples and key results are important and frequently used.

It should be remarked here that scattered throughout Polya's books [17, 18] are hints at some of the elements of this epistemology. For instance, in *Induction and Analogy* [17], he mentions three special kinds of examples --"extreme", "leading", and "representative" -- in several exercises [17, pp. 23-25]. The latter two are very similar to the start-up and model examples discussed in this section. In *Mathematical Discovery*, Polya speaks of the importance of certain "key" facts [18, p. 85]. While he does touch on some of the elements of this epistemology, he does so peripherally to his main points and does not pursue their analysis or use further.

Other authors also single out elements of an epistemology. For instance, Rudin in his classic analysis book *Principles of Mathematical Analysis* uses four headings to organize the presentation: "definitions", "theorems", "examples", and "discussions" [22]. Many authors display some sort of concepts-graph to describe the organization of material, e.g., [21, p.4].

2.4 A Representation Framework

In addition to knowledge of how an item relates to other items, we also have the clusters of information which comprise the item itself. All three of our item types -- results, examples, and concepts -- contain similar pieces of information. For instance, each has a *setting* which is the mathematical context in which the item is known. Each can have a declarative aspect or *statement*: for a concept, a formal mathematical definition; for a result, a (if-then) statement; for an example, a caption, describing what it shows. Each can have a *procedural* aspect: for a concept, a procedural formulation; for a result, a proof; for an example, a construction. An item may have both procedural and declarative aspects (e.g., the eigenvalue concept), or just one (e.g., the Gram-Schmidt process); it may have more than one of either or both aspects (e.g., a result with several proofs). In addition, each item has certain other features such as a worth rating, e.g., the "Michelin" rating [11] which indicates importance by asssignment of from zero to four *'s [3].

We tie these clusters of information together in our representation by amalgamating them into one data structure which has slots for the various component aspects and attributes. All three item types can be represented by the same fundamental framework. This representation framework is then modified slightly for the three item classes of our epistemology -- examples, results, concepts -- to reflect information and features special to them. For instance in the case of results, the representation includes pointers to the converse or more general or stronger results or pointers to counter-examples where these are not possible.

[3] Briefly, the rating scheme is: * for interesting results, worth noticing; ** for important results, worth a "stop"; *** for very important results, worth a "detour"; **** for extremely important results, worth a "journey" in themselves.

Figure 1 shows some of the representation frame for the Cantor set example. Instead of pointers or ID's, we show the name or the statement for item listed in the various pointer fields.

Figure 1

ID *E333* CLASS *Reference, Counter-example* RATING * * * NAME *Cantor Set*

STMNT SETTING *R*

CAPTION *The Cantor set is an example of a perfect, nowhere dense set that has measure zero. It shows that uncountable sets can have measure 0.*

DEMON-STRA-TION AUTHOR *standard*

MAIN-IDEA *Delete "middle-thirds"*

CONSTRUCTION

0. Start with the unit interval [0,1];
1. From [0,1], *delete the middle third* (1/3,2/3);
2. From the two remaining pieces, [0, 1/3] & [2/3, 1], *delete their middle thirds,* (1/9, 2/9) & (7/9, 8/9);
3. From the four remaining pieces, delete the middle thirds;
N. At Nth step, delete from each of the 2^{N-1} *pieces its middle third.*
The sum of the lengths of the pieces removed is 1;
what remains is called the Cantor set.

PICTURE

0__________________ 1

0________( )________ 1

0__( )__( )__( )__ 1

.

.

Limiting set is Cantor set

REMARKS *Cantor set is good for making things happen almost everywhere or almost nowhere.*

LIFTINGS *Construction of general Cantor sets.*

IN-SPACE POINTERS:

BACK *unit interval*

FORWARD *Cantor function, general Cantor sets, 2-dimensional Cantor set*

DUAL-SPACE POINTERS:

CONCEPTS: *countable, measure zero, closed, perfect, geometric series*

RESULTS: *"Perfect sets are uncountable", "Countable sets have measure 0"*

BIBLIOGRAPHIC REFERENCES:

See Gelbaum and Olmstead for details of general Cantor sets.

See Royden for Cantor functions.

PEDAGOGUES *Rudin, Hoffman, Royden*

To organize mathematical knowledge by means of our conceptual framework, we must make several judgments. For instance, recall our arithmetic examples of Section 2. First, we must choose the representation space for an item (e.g., Q, the rational numbers example, could alternatively be classified as a definition), and second, the item must be tied into its chosen space by determining its predecessors and successors (e.g., Q points back to Z, and ahead to R and Q_p). Third, we must link an item to its dual items (e.g., Q can be linked to concepts of division, completeness, density, and cardinality, and to results on the irrationality of $(2)^{1/2}$, and the Archimedian properties of the real line). Fourth, we can sort the dual items into pre- and post-duals. While the specific representation we build reflects certain personal, pedagogical, historical and esthetic biases, the representation scheme is perfectly general.

In summary, our conceptual framework for a mathematical theory includes:

(1) Knowledge of the items themselves: for each we know its statement, diagram, proof, construction or procedural formulation, etc.;

(2) Knowledge of the individual representation spaces and their predecessor-successor relationships;

(3) Knowledge of inter-space relations, such as the dual idea;

(4) Epistemological knowledge of the functional role of items in understanding, such as start-up, reference examples, etc.;

The reader is reminded that this epistemology is neither exhaustive, exclusive nor static. Rather, it represents some important aspects of mathematical knowledge which is a constantly evolving structure. One can view mathematical knowledge as a many-faceted polyhedron that can be held in the hand, rotated, examined from many perspectives, and sliced through along many different planes; our representation tries to capture some of these cross-sectional views, such as its illustrative, pedagogical and inferential aspects.

Also, a particular representation reflects the state of one's knowledge base at a particular moment in time. As long as one keeps learning and thinking, this knowledge base will change and adapt to reflect new knowledge and understanding. Knowledge is not frozen. While it may appear similar for long stretches of one's intellectual time, it is not static. Points of great change or re-organization probably suggest that something important is happening in one's understanding. These aspects are worth looking at further.

3. Understanding as an Active Process

Understanding mathematics is a very active process. While at first glance it may not seem so, especially in comparison with problem solving, it does involve significant effort on the part of the understander. To understand a theory, one must explore and manipulate it on many levels, from many angles, with facility and spontaneity. One must be able to travel freely through it, experiment with its items, survey its overall mathematical topography, shift the level of concern from detail to broad overview and vice versa, and be able to ask questions. One gains understanding by examining relevant examples, perturbing settings and statements, and fiddling around numerically and pictorially. To discover what makes an individual item or a whole theory really work, one must do quite a bit other than passively waiting for understanding to happen.

> "One should try to understand everything: isolated facts by collating them with related facts, the newly discovered through its connection with the already assimilated, the unfamiliar by analogy with the accustomed, special results through generalization, general results by means of suitable specialization, complex situations by dissecting them into their constituent parts, and details by comprehending them within a total picture". [20]

Understanding is a complementary process to problem solving. In many ways it is more difficult to describe than problem solving since, as Polya points out, it is a matter of "more or less and not yes or no" [19]. That is to say, understanding has many levels and is never really totally finished. Actually, understanding, in our sense of building up a knowledge base with all its links and structures, can be taken together with problem solving expertise to comprise a larger view of understanding.

There are many senses and degrees of understanding. Polya abstracts four "levels" of understanding a "rule" from his readings of Spinoza [18, p.134]: (1) "mechanical" when one has memorized the rule and can apply it correctly; (2) "inductive" when one has tried it out in simple cases and is convinced that it works in these cases; (3) "rational" when one has accepted a demonstration of it; and (4) "intuitive" when one is convinced of its truth beyond a doubt.

Poincare also has written about understanding. In particular, he points out the need for going beyond the rational level [15, p.240]:

> "What is it, to understand?...To understand the demonstration of a theorem, is that to examine sucessively each of the syllogisms composing it and to ascertain its correctness, its conformity to the rules of the game? Likewise, to understand a definition, is this merely to recognize that one already knows the meaning of all the terms employed....
>
> For some, yes; when they have done this, they will say: I understand. For the majority, no."

Clearly then, a deep understanding of a theory involves more than knowing just the details of theorems and proofs; it goes beyond simple in-space links. But what should we demand for full understanding? And how should we go about achieving it?

Having deep understanding of a body of mathematics has been likened to knowing one's way around a landscape. Polya and Szego describe it [20]:

> "There is a similarity between knowing one's way about a town and mastering a field of knowledge; from any given point one should be able to reach any other point. One is even better informed if one can immediately take the most convenient and quickest path from the one point to the other. If one is very well informed indeed, one can even execute special feats, for example, to carry out a journey by systematically avoiding certain paths which are customary...
>
> There is an analogy between the task of constructing a well-integrated body of knowledge from acquaintance with isolated truths and the building of a wall out of unhewn stones. One must turn each new insight and each new stone over and over, view it from all sides, attempt to join it on to the edifice at all possible points, until the new finds its suitable place in the already established, in such a way that the areas of contact will be as large as possible and the gaps as small as possible, until the whole forms one firm structure."

Thus if understanding is a matter of "more or less", then clearly deep understanding is a matter of "more". A richness of knowledge is needed for deep understanding.

3.1 Questions that Probe and Prompt Understanding

Despite the lack of widely used, well-defined stages and criteria for understanding, we should not be detered from trying to explicate the understanding process. In this section we offer some questions to help make the process and levels of understanding more crisp and accessible.

When one understands an individual result, concept or example item, one is obviously in command of much information about it. The following questions probe one's understanding of an individual item in the context of a mathematical theory. At the same time, they represent a general strategy for understanding. Being able to answer them is evidence of understanding an item in a thorough way. Being able to ask them indicates knowledge of how to learn.

The intent of this series of questions is not only to make explicit some of the ingredients and processes necessary in the the acquisition of understanding, but also to present them in such a way that a student can learn *how* to go about understanding. Thus the goal is similar to Polya's for problem solving [16] for which his list of "How To Solve It Questions" is offered in the hope of aiding the problem solving process.

The questions are:

1. What is the statement of this item. The setting?

2. Do I understand the statement? Should I review or examine the ingredient concepts, especially the important ones and those to which I have previously not done justice?

3. What is a picture or diagram for this item?

4. Am I reasonably comfortable with this item's immediate predecessors? Are there any predecessors on which I should bone up? Or remember to come back to?

5. Do I know any of dual items for this item, such as counter-examples, model examples, reference examples, culminating results, basic results, etc.? Am I aware of the important ones? Should I peruse some of the others?

6. Can I say what is the gist of this item? Of its statement? Of its demonstration?

7. What is it good for? Why should I bother with it? What is its significance to the theory as a whole?

8. What is the main idea of its proof, construction or procedure? Are the details important? If so, can I summarize them?

9. Is there some way I can fiddle with this item? Perhaps check out a few test cases?

10. What happens if I perturb its statement? Does it generalize? Is it true in other settings? Can it be strengthened by dropping some hypotheses or adding some conclusions. If not, why not: can I cite a counter-example and can I pinpoint what goes wrong? If so, is the new demonstration similar or different from the original. Is it much harder? Should I just be aware that it exists, and forget about the details until I need them?

11. Can I see how this item fits in with the development of the theory as developed in the approach I am taking? What about other approaches? Is this item important or critical or is it simply a stepping-stone or a peripheral embellishment?

12. Can I close my eyes and visualize or describe this item's connections to other items in the theory, to the theory as a whole, to other theories? Have I seen anything like it before?

Clearly this list of questions is rather long and one should not be attempt to answer all of them at once. But one should try to pick off as many questions as possible on an initial try, and if the item is important and worth the effort, come back to the list several times. Through work directly with the item and indirectly with other items, one eventually answers most of the questions. The last question is a keystone to understanding in a deep way and should be given a try during the very first exposure to an item and repreatedly thereafter. At first, the answer given will be very shallow, but later it will become more global and encompassing. It might take two or three passes over the material over several years time perhaps, to be able to expound upon these questions, but that is the fullness of understanding that a mathematician strives for in his work and a student should also set as his goal.

The acquisition of full understanding is often a three pass process. On one's first exposure to a subject, which often occurs while one is taking a course, one tries simply to become familiar with an item and its immediate neighbors (predecessors, successors, pre-dual items). One tries to learn the definitions, read through demonstrations, often checking them out on a step-by-step basis. This first phase is mostly concerned with items one at a time; it is very minimal and local in outlook.

On the second pass, which often comes in reviewing a course, one tries to get a more overall feeling for the subject and the flow of its development. At the least one tries to be able to recall definitions, examples, theorems, and their demonstrations, to see what are the essential assumptions and culminating items, and to know how to get from one item to another. This second phase is concerned with items and relations within the representation spaces and the theory as a whole; it is more global in outlook than the first pass.

The third pass often comes after the course is over, perhaps on another exposure to the material through a different presentation or context, for instance, when listening to a series of lectures "for culture". One starts to see connections between several subjects. One recognizes that the *raison d'etre* of the subject is to address certain questions and that the whole development hinges on certain underlying ideas, axioms or examples; that the subject is very similar to another subject; that many of its items are shared by another subject and are in some sense "the same" as items in another subject. The third pass thus has a perspective that can encompass several theories.

We can correlate these observations and Polya's idea of levels in Spinoza. Our first pass is similar to the "mechanical" and "rational" levels; the second pass, to the "inductive" level; and the third pass, to the "intuitive" level of understanding.

3.2 Knowledge Involved in Understanding

Many of the answers, and the processes needed to find the answers, to these questions can be described in terms of our epistemology. Briefly put, the following information is involved in the answers:

1. the statement and setting of the item;

2. the concepts used in the statement, especially those in the pre-concepts-dual;

3. a picture or diagram for the item;

4. review of predecessor items; tagging of items on bases such as worth or placement on an agenda of items to be examined in future;

5. the item's dual with emphasis on epistemological classes;

6. a paraphrase, synopsis or outline of statement and demonstration.

7. look-ahead through the in-space successor and post-dual items with an eye for important items and epistemological classes.

8. overall structure of demonstration: main idea, plan and skeleton;

9. experimentation with variable elements in statement or picture;

10. perturbation of setting and statement: search and conjecture in more general settings; addition and/or deletion and/or alteration of elements in the statement; look up in references; retrieval of known counter-examples;

11. relations with successors and motivated post-dual items; dependence on predecessors and motivating pre-dual items; knowledge of the (pedagogical) exposition; knowledge of the topography: detours around, direct routes between, and well-worn paths to certain items.

12. intra-space, inter-space, and trans-theory connections; investigation of *sameness* relations through dual and analogy relations.

Thus to understand an item in a deep way, one ought to know about: (1) the item itself; (2) its intra-space relations to other items of the same type; (3) its inter-space relations to other items of different type; (4) dual relations to other items of like type; and (5) relations to items in other theories.

3.3 Understanding A Theory As A Whole

Understanding a theory as a whole is more than just understanding its parts. In addition to understanding member items, it includes understanding the ties that bind the theory together and to other theories. Understanding a theory, like understanding an individual item, involves information about items and connections. In addition, it has a perspective which always seeks to view the item in relation to the whole theory.

Briefly, understanding a theory as a whole involves:

1. knowledge of the epistemological classes: knowing which are the start-up, reference and model examples, the MP's, the CP's, the basic, key and culminating results, etc.: *epistemological knowledge.*

2. knowing the "pros" and "cons" of items: which items are good for what; which items are appropriate and when; how to use them; what their limitations are: *annotative knowledge.*

3. seeing the overall intra-space relations of the individual representation spaces; knowing routes and detours (e.g., "from this item I can get to that one"; "this string of items doesn't lead anywhere"; "the following is a quick and dirty way to derive item X"): knowledge of a *mapping nature.*

4. knowing the inter-space relations such as the items used in recurring dual relations; which items are the basis for striking dual relations; knowing which items are dual equivalent, or nearly so; knowing which items are strikingly similar in the dual sense but are not so within their own representation graphs: knowledge of *sameness* and *closeness*, especially in the sense of the dual idea.

5. abstracting and naming the "arrows", or intra- and inter-space relations, (e.g., Q-->R construction is called "completion" process).

6. recognizing dual and analogy links between items in other theories and theories as a whole: knowledge of *trans-theory links.*

7. recognizing clusters of items generalizing or sharing common features and perhaps eliminating common redundancies and elevating them to the "default", "common sense" or "foundation" knowledge.

4. Classroom Applications

The ideas presented here were used in a seminar with six MIT freshmen. The purpose of this seminar was two-fold: (1) to teach and explore the rich theory of eigenvalues (e.g., the perturbation and location of eigenvalue theorems such as found in Ortega's book [13]); and (2) to make young mathematicians aware of the ingredients and processes involved in understanding mathematics.

The epistemological and organizational ideas seemed natural to the students, especially in discussions in which the students worked out their ideas about keeping track of what they knew and wanted to know. They essentially asked for a representation that included examples, results and definitions, with orderings, and cross-space, i.e., dual, connections.[4] These ideas were also a source of homework problems. For instance, a standard type of problem in the seminar was:

List the dual items for a given item.

Another was:

Tell everything you can about this item.

After the discussion on representation [4] the students were asked as a homework assignment to map out the knowledge domain of the seminar according to our representation scheme; about a month later, they were asked to update their representations. In the seminar we all worked together to meld our representations. While there were some lively debates on how to weave an item into the representation, these sessions always seemed to benefit the students by making them aware of larger issues of how the subject hung together. Thus the organizational process, itself, proved very helpful for developing understanding.

Another type of problem which they enjoyed involved the comparison of theorems addressing a similar topic (e.g., the location of eigenvalues in the Gerschgorin Circle, Symmetric Perturbation, and Hoffman-Wielandt Theorems [13, Chapter 3]):

[4]After about a month, the students wanted to review and catalogue what had thus far been covered in the seminar. At first, they attempted to list all the items in chronological order. Next, they split this list into two lists (definitions and theorems) and then, a third (examples); they tried to order these according to when items occurred. This, they found unsatisfactory since items came up more than once and chronology seemed to have very little to do with anything. Next, they re-ordered results according to what we here have called "logical support", and examples, by a mixture of chronology and increasing complexity; concepts remained in chronological order (which was essentially this author's pedagogical order). This author then told them about directed graphs and trees and with a little prompting, they adopted the three representation graphs of this paper. They were then happily proceeding to organize everything this way in three colors of chalk, when one of the students jumped up, grabbed another color chalk, and pounding his fist on the blackboard, said, "But that's not all there is: each of these results should be connected to some examples and definitions." And so entered the dual idea.

Which theorem is easiest to use, and when?
Which provides the best results, and when?
Cook up at least three (2X2 or 3X3) examples to illustrate your answers.

Most students used reference examples (e.g., the identity, Basic 16) and model examples (e.g., diagonal, upper triangular) in their answers. Together we investigated more complicated matrices with less simple entries (e.g., non-symmetric matrices, matrices with entries of ϵ's and 10^{-1}'s, the Hilbert matrix).

In general, the students displayed a level of mathematical maturity that one would be happy to see in advanced students. They became excellent question askers and idea generators; discussions often left the areas of the author's expertise and entered areas where all were on "hands and knees" together. In short, they became *active.*

4.1 A Theorem Proving Anecdote

Even though the emphasis of this course was not on proving theorems but on understanding them, the following anecdote shows how natural some of the ideas of this paper were to them. One of the students, Ken, requested that we prove the Cayley-Hamilton Theorem (CHT) which states that every matrix A satisfies its own characteristic polynomial, $\det(A-\lambda I)=0$. The students agreed to try to find a proof, but they did not want to work out a purely computational proof involving manipulation of 2x2 and then 3x3 matrices with an induction argument for the general case. Also, we did not want to become involved in considerations of the "minimal polynomial" and its attendant algebra. The following is a nearly verbatim report of the dialogue that ensued when the students were asked to suggest a plan of attack:

JOHN *The theorem is certainly true for the identity matrix.*

DAVID *Check. Further if the CHT is true in general, it must be true for diagonal matrices. Right?*

ERM *Right.*

JOHN *That case is easy.*

DAVID *OK. So now we should be able to show it's true for diagonalizable matrices, by using the similarity transform $S^{-1}DS$, on diagonal matrices and hoping that the algebra goes away.*

KEN *So?*

DAVID *So, then maybe we can get the general case by doing the same thing on upper triangular matrices and using the fact, i.e., the Jordan Normal Form Theorem, which we haven't proved, but know about and all believe, that all matrices are similar to upper triangular matrices with their eigenvalues on the main diagonal.*

KEN *That sounds good to me.*

JOHN *Does all the algebra come out right?*

ERM *Let's try it and see.*

And so we developed David's plan by establishing the theorem in the upper triangular case ([25, p. 224] gives an approach to this) and it did indeed lead to a proof of the theorem. There are several noteworthy features about this episode: (1) the line of reasoning parallels *exactly* the direction of constructional derivation of one branch of the examples-graph we built: Identity --> Diagonal --> Upper Triangular; (2) they strongly used reference and model examples (e.g., identity, diagonal and upper triangular matrices) of the eigenanalysis domain; (3) the whole interchange was completely spontaneous and took but a minute. The rest of the seminar was truly amazed at the speed at which David formulated his plan, and also how pretty it was. David commented that it seemed the "obvious" thing to do. Ken chose to write about this theorem, its proof and the importance of examples as his term paper.

4.2 Some Comments on Problem Solving

During the semester, the students met to work on some selected problems in a one-on-one manner. The ground rules were that: these sessions were not tests; they could look up anything they wanted in our notes and references; they could always ask for suggestions and advice; there were no time constraints; and if possible, they would try to think out loud while they worked.

All the sessions were tape-recorded. The problems ranged in difficulty and style from standard questions with a stated goal, such as:

Show that the possible eigenvalues of an involution ($U^2=I$) are +1 and -1.

or:

Give a counter-example to show that interchanging rows of a matrix does not leave its eigenvalues unchanged.

to more vaguely-posed problems, such as:

What can you say about the spectrum of a permutation matrix?

Most all of the students handled the first question by using the declarative definition for "eigenvalue". All the students answered the second question by examining the reference collection of the "Basic 16". Most attacked the third question by examining the 2X2 cases to form a preliminary conjecture and then some of the 3X3 cases to test and refine it; not all started out this way, but those that tried to attack the problem through more general arguments found they could not get a handle on the problem and thus followed the heuristic of examining the two-dimensional case. To this author's delight they handled these problems with great poise and enthusiasm. They were, for the most part, completely undaunted by the fact that they had to decide how to attack the open-ended problems. As a bonus their answers were very complete.

5. Understanding Mathematics

Understanding mathematics is a process that can be understood and to some extent taught. In our view of understanding, a good part of the process is concerned with building and enriching a knowledge base. This includes creating associations of many kinds as well as items. It also involves differentiating between various kinds of items according to their function in acquiring knowledge, familiarity, and expertise.

In summary, some of the ingredients of the process of understanding mathematics are:

1. Knowledge of items and relations: general types such as the item/relation pairs of the three representation spaces and dual relations, as well as particular ones such as generalization and specialization;

2. General strategic or control knowledge such as: knowing to restrict the situation under consideration to the particular case of an example, such as a reference example; in particular, restricting the situation under consideration to the case of an example of known generality, such as a model example, analysing how things work, and then lifting back up; knowing to fool around with examples, especially reference or models, when out of ideas; knowing to perturb statements and settings;

3. Meta-knowledge such as: knowing to keep one's eyes open for items of special note such as models, references, MP's, etc.; and knowing that keeping track of links by mapping out one's knowledge base (at least thinking about trying to do this) can be a useful not only to keep track of what one knows but to build global understanding;

4. Epistemological knowledge : knowing that certain items serve particular

functions in understanding; and that some ideas and processes, such as the "group" idea [1] or the "divide and conquer" technique [11] are very general and pervasive through all of mathematics.

5. <u>Representational knowledge</u> of knowing how to organize and keep track of what one knows such as through maps and networks of items and relations, and through representation schemes, such as frameworks for individual items.

Thus, to understand an item or a theory fully, one must be able to examine it at different levels of detail and from several points of view; follow infra-space and inter-space associations; perturb and fiddle with items; and survey the overall topography of the spaces individually and together; and link them with other theories. In short, to achieve a deep sense of understanding one must have established many links of all kinds.

References

[1] Bourbaki, N., "The Architecture of Mathematics". The American Mathematical Monthly, Vol 57 (1950), pp. 221-232.

[2] Eilenberg, S., and N. Steenrod, *Foundations of Algebraic Topology*. Princeton University Press, New Jersey, 1973

[3] Freudenthal, H., *Mathematics as an Educational Task*. Reidel, Holland, 1973.

[4] Gelbaum, B. R., and J. Olmstead, *Counterexamples in Analysis*. Holden-Day, California, 1964.

[5] Goldstein, I. P., "The Genetic Graph". Proceedings Second National Conference of Canadian Society for Computational Studies of Intelligence, 1978.

[6] Halmos, P. R., *Finite-Dimensional Vector Spaces*. D. Van Nostrand, New Jersey, 1942.

[7] Hoffman, K. M., *Analysis in Euclidean Space*. Prentice-Hall, New Jersey, 1975.

[8] Ireland, K., and M. I. Rosen, *Elements of Number Theory*. Bogden and Quigley, New York, 1972.

[9] Jans, J. P., *Rings and Homology*. Holt, Rinehart and Winston, New York, 1964.

[10] Lakatos, I., *Proofs and Refutations*, British Journal for the Philosophy of Science, Vol 19, May 1963, No. 3. Also published by Cambridge University Press, London, 1976.

[11] Michener, E. R., *Epistemology, Representation, Understanding and Interactive Exploration of Mathematical Theories*. Doctoral dissertation MIT Department of Mathematics, 1977.

[12] Michener, E. R., *The Structure of Mathematical Knowledge*. Technical Report, MIT Artificial Intelligence Laboratory, forthcoming.

[13] Ortega, J. M., *Numerical Analysis: A Second Course*. Academic Press, New York, 1972.

[14] Piaget, J. *Genetic Epistemology*. W. W. Norton and Company, New York, 1970.

[15] Poincare, H., *Foundations of Science*. The Science Press, New York, 1929.

[16] Polya, G., *How To Solve It*. Second Edition, Princeton University Press, New Jersey, 1973.

[17] _______, *Induction and Analogy in Mathematics*, Volume 1 of *Mathematics and Plausible Reasoning*. Princeton University Press, New Jersey, 1973.

[18] _______, *Mathematical Discovery*, Volume 2. Wiley, New York, 1962.

[19] _______, *Personal communication*. 1978.

[20] Polya, G., and G. Szego, *Problems and Theorems in Analysis I*. Springer-Verlag, New York, 1972.

[21] Royden H. L., *Real Analysis*. Second edition, Macmillan, New York, 1963.

[22] Rudin, W., *Principles of Mathematical Analysis*. Second Edition, McGraw-Hill, New York, 1964.

[23] Schoenfeld, A. H., *Can Heuristics Be Taught*. SESAME Project, University of California at Berkeley, 1978.

[24] Spivak, M., *Differential Geometry*. Vol. 2, Publish or Perish Press, Boston, 1972.

[25] Strang, G., *Linear Algebra and its Applications*. Academic Press, New York, 1976.

[26] Thomas, G. B., *Calculus and Analytic Geometry*. Fourth Edition, Addison-Wesley, Massachusetts, 1972.