DOCUMENT RESUME

ED 201 505 SE 034 846

AUTHOR Schoenfeld, Alan H.
TITLE Episodes and Executive Decisions in Mathematical Problem Solving.
PUB DATE Apr 81
NOTE 73p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981). Contains occasional light and broken type.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Cognitive Processes; *College Mathematics; Educational Research; Evaluation; *Geometric Concepts; Geometry; Higher Education; Learning Problems; *Learning Theories; *Mathematics Education; Mathematics Instruction; *Problem Solving; Undergraduate Study
IDENTIFIERS *Mathematics Education Research

ABSTRACT

The research described here seeks to characterize the "managerial" aspects of expert and novice problem-solving behavior, and to describe the impact of managerial or "executive" actions on success or failure in problem solving. A framework for analyzing protocols of problem-solving sessions based on "episodes" of problem-solving behavior and focusing on managerial decisions between episodes is presented. Experts are shown to have rather "vigilant" managers, which strive for efficiency and accuracy. In contrast, novices squander their problem-solving resources because they lack such managers. (Author/MP)

EPISODES AND EXECUTIVE DECISIONS IN MATHEMATICAL PROBLEM SOLVING*

Alan H. Schoenfeld
Mathematics Department
Hamilton College
Clinton, N.Y. 1332[illegible]

*This paper was presented at the [illegible] AERA Annual Meeting.

Running Head: Episodes and Executive Decisions

EPISODES AND EXECUTIVE DECISIONS IN MATHEMATICAL PROBLEM SOLVING*

Alan H. Schoenfeld
Mathematics Department
Hamilton College
Clinton, N.Y. 13323

*This paper was presented at the 1981 AERA Annual Meeting.

Running Head: Episodes and Executive Decisions

Abstract

The research described here seeks to characterize the "managerial" aspects of expert and novice problem-solving behavior, and to describe the impact of managerial or "executive" actions on success or failure in problem solving. We present a framework for analyzing protocols of problem-solving sessions based on "episodes" of problem-solving behavior and focusing on managerial decisions between episodes. Experts are shown to have rather "vigilant" managers, which strive for efficiency and accuracy. In contrast, novices squander their problem-solving resources because they lack such managers.

Episodes and Executive Decisions in Mathematical Problem Solving

Introduction and Overview

This is a rather speculative paper dealing with "managerial" decisions in human problem solving. It presents a (still evolving) framework for the analysis at the macroscopic level of problem-solving protocols, focusing on "executive" behaviors. The paper is based on the following premise.

There are two qualitatively different kinds of decisions, which we shall call "tactical" and "strategic," which are necessary in broad, semantically rich domains (for example, mathematical problem solving at the college freshman level). The first, tactical decision making, has received the lion's share of attention. By tactics I mean "things to implement." Tactics include all algorithms and most heuristics, both of the Pólya type (e.g., draw a diagram whenever possible; consider special cases) and of the kinds used in Artificial Intelligence (means-ends analysis, hill-climbing). Given that one has decided to calculate the area of a particular region, the choice of whether to approach that calculation via trigonometry or analytic geometry is a tactical choice.

In contrast, "strategic" or managerial decisions are those which have a major impact on the direction a solution will take, and on the allocation of one's resources during the problem-solving process. For example: If one is given twenty minutes to work on a problem and calculating the area of a region is likely to take ten minutes, the decision to calculate the area of that region is a strategic one--regardless of the method ultimately chosen for performing the calculations. Like a decision during wartime to open a front, this one choice may determine the success or failure of the entire enterprise.

This separation of managerial decisions from implementation decisions has implications for both human and machine problem solving. Mathematics problem-solving instruction to date has focused largely, and with somewhat questionable success, on heuristics or "tactics." I propose that much of the reason for this lack of success lies in the fact that attention to managerial behaviors has mostly been neglected. The protocols discussed below will indicate that heuristic fluency is of little value if the heuristics are not "managed" properly. I believe that much greater attention will have to be paid to "metaheuristics" or managerial actions in classroom instruction, if we are to be successful in teaching problem-solving skills.

There appear to be parallels in artificial intelligence. Regardless of their sophistication, production systems are essentially tactical decision-makers. They are not strategists. The managerial decisions made in such programs, by "conflict resolution strategies" when the conditions for more than one production are met simultaneously, seem to be more or less <u>ad hoc</u> and idiosyncratic, rather than theory-based. For the most part, programming in narrow domains finesses the question of managerial strategies. However, such concerns cannot be ignored as the domains of investigation are broadened. Further, some attempt at dealing with executive strategies must be made for the creation of "glass box" experts in computer-based tutorial systems for non-trivial domains. Since such decisions are an important component of human problem solving, any system in a broad arena which ignores them will lack psychological validity.

This paper discusses a framework for examining, at the macroscopic level, a broad spectrum of problem-solving protocols. Protocols are parsed into major "episodes." These are periods of time during which the problem solver(s) is engaged on a single set of like actions, such as "planning" or "exploration." It is precisely between such episodes that the managerial decisions which can "make or break" a solution are often made, or not made. We focus on decision making at these points, and on the impact of such decisions--or their absence--on problem-solving performance. The quality and success of problem-solving endeavors will seem to correspond closely (in human problem solving) to the presence, and vigilance, of such "managers."

A Discussion of Antecedents

By definition, protocol coding schemes are concerned with producing objective records or "traces" of a sequence of overt actions taken by individuals in the process of solving problems. In mathematics education, the coded protocol is generally subjected to a qualitative analysis; often correlations will be sought between certain types of behavior (e.g., the presence of goal-oriented heuristics) and problem-solving success. In artificial intelligence, the goal is often to write a program that will simulate a given protocol, or the idealized behavior culled from a variety of protocols. In both cases the level of analysis is microscopic. My goal here is to indicate that in many cases the microscopic level analysis may be entirely inappropriate. In analyzing human problem solving, attention to that level of detail may cause one to "miss the forest for the trees"; if the wrong strategic decisions are made, tactical ones are virtually irrelevant. In artificial intelligence, great progress has been made at the tactical level through the use of production systems. It is not at all clear, however, that they will serve well for making managerial decisions. I believe that we may wish to think of these executive decisions as being at a higher level than tactical ones, and may want to deal with these "strategists" separately.

(Note: what follows is an opinionated discussion of the recent literature, which depends heavily on the distinction between "tactical" or "strategic" or "managerial" decisions. These distinctions may be much clearer after the reader has considered the examples discussed in the next section. Thus the reader may wish to skip ahead to that section, and later consider the comments made here in the light of those examples.)

The following description, taken from Lucas et al., (1979, p. 354) is typical of the efforts of mathematics educators to deal with problem-solving protocols.

> [T]he authors came to agreement on the definitions for a set of constructs which were to represent observable, disjoint problem solving behaviors and related phenomena . . . . Each event was assigned a symbol, and the collection of events which comprised a problem-solving sequence of processes was recorded in a horizontal string of symbols corresponding to the chronological order of appearance during the actual problem solution. In this manner a researcher could listen to a tape of a problem solution (in conjunction with observing written work, interviewer notes, and/or a verbatim transcript) and produce a string of symbols which represented the composite perception of the solution process. Conversely, an examination of the given string of symbols could be used to provide a reasonably clear picture of what had happened during a problem-solving episode.

That particular coding scheme included a two-page "dictionary" of processes which were assigned coding symbols. All behavior was "required to be explicit; otherwise it is not coded." (p. 359) As an example of the coding, the sequence (p. 361)

> The problem solver reads the problem, hesitates, rereads part of the problem, says the problem resembles another problem and he will try to use the same method, then deduces correctly

a piece of information from one of the given data was coded as $(R,R,L_8P_iD_{a5}4)$.

In part because of the cumbersome nature of such systems and the wealth of symbols that must be dealt with, once coded, other researchers have opted to focus on more restricted subsets of behaviors. Kulm's recent NSF-supported work, "Analysis and Synthesis of Mathematical Problem Solving Processes," uses a revised and more condensed process code dictionary (private communication, 1979). Kantowski's recent work (Note 3) includes a "coding scheme for heuristic processes of interest" which focuses on five heuristic processes related to planning, four related to memory for similar problems, and seven related to looking back. The frequency of such processes is related to problem-solving performance.

So far as I know, there are no systems for protocol analysis that focus in any substantive way on strategic decisions. There are no frameworks for dealing with things which ought to have been considered, but were not. For the most part, discussions in the literature of executive decision making during problem solving are weak. Polya, for example (1965, p. 96) offers "Rules of Preference" for choosing among options in a problem-solving task. These include injunctions such as "the less difficult precedes the more difficult" and "Formerly solved problems having the same kind of unknown as the present problem precede other formerly solved problems." My own attempts (Schoenfeld, 1979; 1980) at capturing a managerial strategy in flow chart form for students' implementation were somewhat impoverished, the flow chart in effect presenting a default strategy. All other factors being equal--meaning that the problem solver had exhausted the lines of attack which had appeared fruitful (his "productions?") and had no strong leads to follow up--it was considered reasonable to try the heuristic suggestions in this "managerial strategy," roughly in the manner suggested by the flow chart. This bypassed the tough questions, however. Issues like: how does one decide what to pursue; for how long; how does one evaluate progress towards a solution; when should the "manager" interfere, etc.,while discussed in class, were not formally a part of the strategy. Moreover, there was no systematic and rigorous framework for examining these questions.

As a result of (1) the narrowness of the problem domains in which artificial intelligence has successfully operated, and (2) the tactical utility of production systems in those domains, the AI community has given even less attention to executive strategies than has the math-ed community. The questions are not new: the "considerations at a position in problem space" listed by Allen Newell (1966, figure 5) are quite similar to those we will pose below. But

> "Select new operator:
>
> Has it been used before?
>
> Is it desirable: will it lead to progress?
>
> Is it feasible: will it work in the present situation if applied?"

takes on very different shades of meaning at the strategic rather than the tactical level. So far as I can tell, (and my knowledge of such is limited) recent advances in production systems allow for rather clever tactical decision making. There are computationally efficient means of keeping track of and sorting through productions for relevancy, and there are conflict resolutions systems (McDermott & Forgy, 1978) for selecting among productions when the conditions for more than one of them have been satisfied. Such structures prohibit productions from executing more than once on the same data. This prevents the kind of endless reptitions all too common in students and forces, if necessary, the examination of all available information. Since preference is given to productions whose conditions are satisfied by elements most recently placed in working memory, there is a "natural" continuity to the sequence of operations. Other means of selection (e.g., specificity precedes generality) provide plausible means of selecting tactics in relatively narrow domains. Yet I am not sure that the level of analysis is right for general problem solving, or that such strategies would have much to say about the strategic decisions in the examples given in the next section. Similar comments apply to the "adaptive" or "self-modifying" production systems described by Anzai and Simon (1979), Neves (1978), and Neches (1979). While the learning principles they exemplify may be general, the embodiments of those principles in those papers are at the tactical level. Simon (1980) argues that "effective professional education calls for attention to both subject matter knowledge and general skills (p. 86)" and then goes on to say (p. 91) that "general skills (e.g., means-ends analysis) will be particularly important in the learning stages but will also show up implicitly in the <u>form</u> of the productions that are used in the skilled performance." But even this is one step removed from the heart of the matter: what underlies the <u>form</u> of the productions is in the mind of the programmer, not in the productions. We need a methodology for focusing on those general skills directly.

An Informal Analysis of Two Protocols

The AI literature is filled with beautiful protocols. I have never been that lucky: those generated by my students (and to some extent by my colleagues) in the process of grappling with relatively unfamiliar problems have been, on the whole, rather unaesthetic. This section considers two such protocols, each generated by a pair of students. (Following a suggestion from John Seely Brown, I have students work on problems in pairs. While the question "why did you do that?" coming from me may be terribly intimidating and is likely to alter the solution path, the question "why should we do that?" from a fellow student working on a problem is not. This type of dialogue between students often serves to make managerial decisions overt, whereas such decisions are rarely overt in single-student protocols.) An informal analysis, focusing on the importance of managerial decisions, follows.The formal analytic structure is given in the next section.

Protocols 1 and 2 are given in Appendices 1 and 2, respectively. The students were asked to work on the problem together, out loud, as a collaborative effort. They were not to go out of their way to explain things for the tape, if that interfered with their problem solving; their interactions, if truly collaborative, would provide me with the information I needed. (See Ericsson and Simon (1978; 1979) for a discussion of instructions for speak-aloud experiments.) All of the students were undergraduates at a liberal arts college. Students A and K (protocol 1) had 3 and 1 semesters of college mathematics (calculus) respectively. Students D and B (protocol 2) each had 3 semesters of college mathematics. It should be recalled that such students, by most standards, are successful problem solvers: the unsuccessful ones had long since stopped taking mathematics courses. Both protocols are of the same problem:

> Three points are chosen in the circumference of a circle of radius R, and the triangle containing them is drawn. What choice of points results in the triangle with the largest possible area? Justify your answer as best you can.

If protocol 1 makes for confused reading, the tape it was taken from makes for even more pained viewing. I would summarize the problem-solving session as follows:

The students read and understood the problem, and then quickly conjectured that the answer was the equilateral triangle. They impetuously decided to calculate the area of the triangle, and spent the next 20 minutes doing so. These calculations of the area were occasionally punctuated by suggestions which might have salvaged the solution, but in each case the suggestions were quickly dropped and the students returned to their relentless pursuit of the worthless calculation. (Neither student could tell me, after the cassette ran out of tape, what good it would do them to know the area of the equilateral triangle.) Observe the following.

1. The single most important event in the twenty-minute problem-solving session, upon which the success or failure of the entire endeavor rested, was one which did not take place--the students did not assess the potential utility of their planned actions, calculating the area of the equilateral triangle. In consequence, the entire session was spent on a wild goose chase.
2. Inadequate consideration was given to the utility of potential alternatives which arose (and then submerged) during the problem-solving process. Any of these: the related problem of maximizing a rectangle in a circle (item

28), the potential application of the calculus (item 52) for what can indeed be considered a max-min problem; the qualitative varying of triangle shape (item 68) might have, if pursued, led to progress. Instead, the alternatives simply faded out of the picture. (See, for example items 27 to 31.)

3. Progress is never monitored or (re)assessed, so that there is no reliable means of terminating wild goose chases once they have begun. (This is to be strongly contrasted with an expert protocol, where the problem solver interrupted the implementation of an outlined solution with "this is too complicated. I know the problem shouldn't be this hard.")

Now, how does one code such a protocol? First, we should observe that matters of detail (such as whether or not the students will accurately remember the formula for the area of an equilateral triangle, items 73 to 75) are virtually irrelevant. To return to the military analogy in the opening section: if it was a major strategic mistake to open a second front in a war, the details of how a hill was taken in a minor skirmish on that front are of marginal interest.

A second and more crucial point is that the overt actions taken by the problem solvers in that protocol are, in a sense, of minor import. The problem-solving effort was a failure because of the absence of assessments and strategic decisions. Any framework that will make sense of that protocol must go beyond simply recording what did happen; it should suggest when strategic decisions ought to have been made, and allow one to interpret success or failure in the light of whether, and how well, such decisions were made.

If protocol 1 stands as evidence of the damage that can be caused by a manager "in absentia," protocol 2 provides evidence of the catastrophic effects of bad management. The processes in this tape were not muddled, as in protocol 1; the decisions were overt and clear. The next paragraph summarizes the essential occurrences in the tape. The superscripts refer to the commentary that follows.

D and B quickly conjecture that the solution is the equilateral triangle, and look for ways to show it. D, apparently wishing to exploit symmetry in some way, suggests that they examine triangles in a semicircle with one side as diameter. They find the optimum under these constraints, and reject it "by eye" as inferior to the equilateral.[1] Still focusing on symmetry, they decide[2] to maximize the area of a right triangle in a semicircle, where the right angle lies on the diameter. This (serendipitously correct) decision reduces the original problem to a 1-variable calculus problem[3] which B proceeds to work on. Twelve minutes later the attempt is abandoned,[4] and the solution process degenerates into an aimless series of explorations, most of which serve to rehash the previous work.[5]

1. Rejecting the alternative is quite reasonable, as are their actions in analyzing the problem up to this point. However, this blanket rejection may have cost them a great deal. The variational argument they used to find the isosceles right triangle (holding the base fixed and observing that the area is largest when the triangle is isosceles) is perfectly general and can be used to solve the original problem as stated. But the students simply turn away from their unsuccessful attempt, without asking if they could learn from it. In doing so, they may have "thrown out the baby with the bath water."

2. This decision, which affects the direction of the solution for more than 60% of the allotted time, is made in a remarkably casual way (items 24 to 27):

   D: (after one attempt at symmetry has failed) ...you want to make it perfectly symmetrical, but we can, if we maximize this area, just flip it over, if we assume that it is going to be symmetrical.

   B: Yea, it is symmetrical.

   This assumption is not at all justified (they are assuming part of what they are to prove). The students have changed the problem and proceed, without apparent concern, to work on the altered version.

3. B's tactical work here is quite decent, as is much of both students' tactical work throughout the solution process. The decision to "scale down" the problem to the unit circle (item 37) is just one example of their proficiency. There is awareness of, and access to, a variety of heuristics and algorithmic techniques during the solution. Unfortunately, B lost a minus sign during this particular calculation, which gave him a physically impossible answer. He was aware of it; local assessment worked well. However, global assessment (see 4 and 5) did not.

4. This decision to abandon the analytic approach is just as astonishing, in the way it takes place (items 74 and 75) as the decision to undertake it:

   D: Well, let's leave the numbers for a while and see if we can do it geometrically.

   B: Yea, you're probably right.

   Given that more than 60% of the solution has been devoted to that approach (and that correcting a minor mistake would salvage the entire operation), this casual dismissal of their previous efforts has rather serious consequences.

5. There were a number of clever ideas in the earlier attempts made by D and B. Had there been an attempt at a careful review of those attempts, something might have been salvaged. Instead, there was simply a "once over lightly" of the previous work that added nothing to what they had already done.

A framework for focusing on the managerial decisions in such protocols is discussed in the next section.

A (poorly defined and still evolving) Framework for the Macroscopic Analysis of Certain Kinds of Problem-Solving Protocols

The two protocols discussed in the preceding section raise the major questions I wish to address here. I believe that decisions at the managerial level may "make or break" a problem-solving attempt, and that (at least in the case of poor managerial decisions) these may render irrelevant any subsequent tactical (i.e., implementation) decisions. Thus we focus on behavior at the macroscopic level.

Protocol 1, which is rather typical of students' problem solving, illustrates one of the major difficulties in dealing with managerial decisions: the absence of intelligent management may doom problem-solving attempts to failure. Yet all extant schemes focus on what is overtly present, ignoring the crucial decisions that might (and should!) have taken place. Protocol 2 is, in a sense, easier to deal with. The decisions were overt, though poor. This protocol serves to indicate that decision making means more than simply choosing solution paths: it incorporates local and global assessments of progress, as well as trying to salvage the valuable elements of ultimately flawed approaches. This section offers a scheme for parsing protocols that tries to address these issues.

There are both objective and subjective components to the framework for analyzing protocols. The objective part consists of identifying, in the protocol, the loci of potential managerial decisions. The subjective part consists of characterizing the nature of the decision-making process at these "managerial decision points" and describing the impact of those decisions (or their absence!) on the overall problem-solving process.

By definition, managerial or strategic action is appropriate whenever a large amount of tactical resources are about to be expended. This provides the basic idea for parsing the protocols. Partition a protocol into macroscopic chunks of consistent behavior ("episodes"). Then the points between episodes--where the direction or nature of the problem solution changes significantly--are the managerial decision points where, at minimum, managerial action ought to have been considered.

In addition to these junctures between episodes, there are two other loci for managerial action: at the arrival of new information or the suggestion of new tactics, and at the point where a series of tactical failures indicates that strategic review might be appropriate. The loci that deal with new information are well defined and pose little difficulty in identification. Observe that this kind of decision point can occur in the middle of an episode: new information may be ignored or dismissed (at least temporarily), and the problem solver may continue working along previously established lines. The latter kind is more difficult, and calls for subjective judgment; I have no easy way of dealing with these at present. At some point when implementation bogs down, or when the problem-solving process degenerates into more or less unstructured explorations, it is time for an "executive review." It is clear from the protocols I have taken that experts have "monitors" that call for such review, and that novices often lack them. We will return to this point later, in the subjective analysis.

Figures 1 and 2 represent a parsing of protocols 1 and 2, respectively, into episodes. "New information" points within episodes are indicated.

Insert Figures 1 and 2 about here

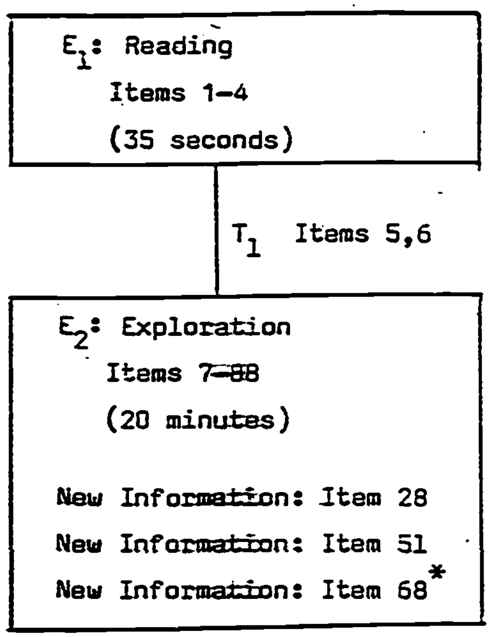


Figure 1

A Parsing of Protocol 1

*Note: From the written protocol it might appear that Item 68 begins a new episode. In fact, the students had lost virtually all their energy by that point, and were merely doodling; they returned (after the tape clicked off) to musings about the equilateral triangle. Thus items 6-88 are considered to be one episode.

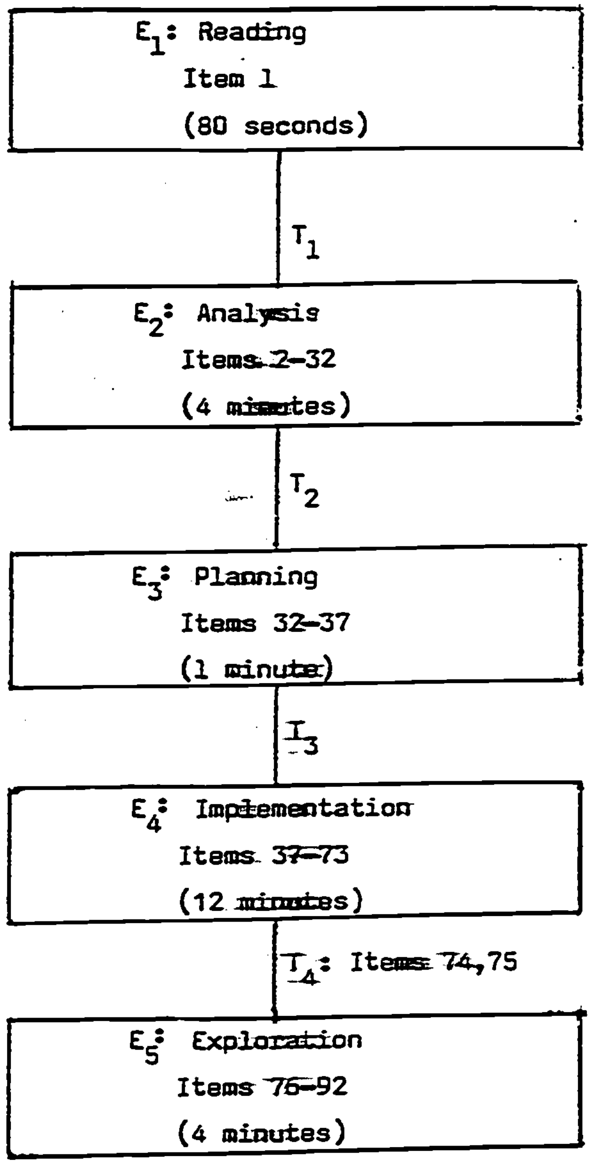


Figure 2

A Parsing of Protocol 2

Detailed analyses of Figures 1 and 2 will not be given, since protocols 1 and 2 have been discussed at some length. (Observe, however, how Figures 1 and 2 reflect the issues singled out for discussion above.) A third protocol will be analyzed in detail.

Both parsing into episodes and delineating "new information" points, turn out to be (more or less) objective decisions. In fact, the parsing of all three protocols that I use in this paper was derived, in consensus, by three undergraduates who followed my instructions but arrived at their characterizations of the protocols in my absence. Reliability in parsing protocols is quite high. (This does not, however, obviate the need for an appropriate formalism: see the final commentary.)

Subjectivity lurks around the corner, however. It is, in fact, already present in the labeling of the episodes given in Figures 1 and 2. This labeling was essential: see the note below.* Any episode is characterized as one

*The potential for "combinatorial explosion" in characterizing managerial behaviors is enormous. Managerial behaviors include selecting perspectives and frameworks for a problem; deciding at branch points which direction a solution should take; deciding whether, in the light of new information, a path already embarked upon should be abandoned; deciding what (if anything) should be salvaged from attempts that are abandoned or paths that are not taken; monitoring tactical implementation against a template of expectations for signs that intervention might be appropriate; and much, much more. My early attempts at analyses of managerial behavior called for examining protocols at all managerial decision points and evaluating at each one a series of questions encompassing the issues just mentioned. This approach, while comprehensive, was completely unwieldy. For example, questions about the assessment of state when (a) one has just read the problem, (b) one is "stuck," and (c) a solution has been obtained, are almost mutually exclusive. Thus at any decision point 90% of the questions that might be asked were irrelevant. The framework described above provides a workable compromise.

of the following: Reading, Analysis, Planning, Implementation (or Planning/Implementation if the two are linked), Exploration, Verification, or Transition. What follows is the heart of the analytic framework. There is a brief description of the nature of each type of episode, followed by a series of questions to be asked about each episode once it has been labeled. The parsing, plus the answers to the questions, provide the characterization of the protocol.

Admittedly, these questions are a mixed bag. Some can be answered objectively at the point in the protocol at which they are asked, some in the light of later evidence; some call for inferences or judgments about problem-solving behavior. Further, some ask about the "reasonableness" of certain behavior. Asking questions in this way, of course, begs the significant question: what is a model of "reasonable" behavior? The creation of such models is the crucial long-term question, and there is no attempt to finesse it here. At present, however, we will deal with the notion subjectively, to better understand managerial behaviors so that we can create those models. Though highly subjective, these assessments can be made reliably: agreement between my ratings and the consensus scorings of my students was quite high. To quote Mr. Justice Stewart (1964), "I shall not today attempt to further define the kind of materials I understand to be embraced within that shorthand definition;... But I know it when I see it."

Episodes and the Associated Questions

1. READING.

The reading episode begins when a subject starts to read the problem statement aloud. It includes the ingestion of the problem conditions, and continues through any silence that may follow the reading--silence that may indicate contemplation of the problem statement, the (non-vocal) rereading of the problem, or blank thoughts. It continues as well through vocal re-readings and verbalizations of parts of the problem statement (observe that in protocol 1, reading included items 1-4).

READING Questions:

a. Have all of the conditions of the problem been noted? (Explicitly or implicitly?)

b. Has the goal state been correctly noted? (Again, explicitly or implicitly?)

c. Is there an assessment of the current state of the problem solver's knowledge relative to the problem-solving task (see TRANSITION)?

2. ANALYSIS.

If there is no apparent way to proceed after the problem has been read (i.e., a solution is not "schema driven"), the next (ideal) phase of a problem solution is analysis. In analysis, an attempt is made to fully understand a problem, to select an appropriate perspective and to reformulate the problem in those terms, and to introduce for consideration whatever principles or mechanisms might be appropriate. The problem may be simplified or reformulated. (Often analysis leads directly into plan development, in which case it serves as a transition. Of course, this episode may be bypassed completely.)

ANALYSIS questions:

a. What choice of perspective is made? Is the choice made explicitly, or by default?

b. Are the actions driven by the conditions of the problem? (working forwards)

c. Are the actions driven by the goals of the problem? (working backwards)

d. Is a relationship between conditions and goals sought?

e. Is the episode, as a whole, coherent? In sum (considering a-d), are the actions reasonable? (comments?)

3. EXPLORATION.

Both its structure and content serve to distinguish exploration from analysis. Analysis is generally well structured, sticking rather closely to the conditions or goals of the problem. Exploration, on the other hand, is less well structured and further removed from the original problem. It is a broad tour through the problem space, a search for relevant information that can be incorporated into the analysis/plan/implementation sequence. (One may well return to analysis with new information gleaned during exploration.)

In the exploration phase of problem solving one may find a variety of problem-solving heuristics, the examination of related problems, the use of analogies, etc. Though amorphously structured, exploration is not, ideally, without structure: there is a loose metric on the problem space, the perceived distance of objects under consideration from the original problem, that should serve to select items for consideration. Precisely because exploration is weakly structured, both local and global assessments are critical here (see transition as well). A wild goose chase, unchecked, can lead to disaster; but so can the dismissal of a promising alternative.

If new information arises during exploration but is not used, or the examination of it is tentative, "fading in and fading out," the coding scheme calls for delineating "new information" within the episode. If, however, the problem solver decides to abandon one approach and start another, the coding scheme calls for closing the first episode, denoting (and examining) the transition, and opening another exploration episode.

EXPLORATION questions:

a. Is the episode condition driven? Goal driven?

b. Is the action directed or focused? Is it purposeful?

c. Is there any monitoring of progress? What are the consequences for the solution of the presence or absence of such monitoring?

d. At NEW INFORMATION points (including the introduction of heuristics) and LOCAL ASSESSMENT points:

1. Does the problem solver assess the current state of his knowledge? (Was it appropriate??)

2. Does the problem solver assess the relevancy or utility of the new information? (Was it appropriate?)

3. What are the consequences for the solution of the actions (or inactions) described in 1 and 2 above?

4. PLANNING/IMPLEMENTATION.

Since the emphasis here is on managerial questions, detailed issues regarding plan formation will not be addressed: the primary questions of concern here deal with whether or not the plan is well-structured, whether the implementation of the plan is orderly, and whether there is monitoring or assessment of the process on the part of the problem solver(s), with feedback to planning and assessment at local and/or global levels. Many of these judgments are subjective. For example, the absence of any overt planning behavior does not necessarily indicate the absence of a plan; in fact, protocols of "schema-driven" solutions often proceed directly from the reading episode into the coherent and well structured implementation of a non-verbalized plan. Thus the latitude of the questions below: the scheme should apply to a range of circumstances, from schema-driven solutions to those where the subject happens upon an appropriate plan by design or accident.

PLANNING/IMPLEMENTATION questions:

a. Is there evidence of planning at all? Is the planning overt or must the presence of a plan be inferred from the purposefulness of the subject's behavior?

b. Is the plan relevant to the problem solution? Is it appropriate? Is it well structured?

c. Does the subject assess the quality of the plan as to relevance, appropriateness, or structure? (If so, how do those assessments compare with the judgments in (b)?)

d. Does implementation follow the plan in a structured way?

e. Is there assessment of implementation (especially if things go wrong), at the local or global level?

f. What are the consequences for the solution of assessments if they occur, or if they do not?

5. VERIFICATION.

The nature of the episode itself is obvious.

a. Does the problem solver review the solution?

b. Is the solution tested in any way? (If so, how?)

c. Is there any assessment of the solution, either an evaluation of the process or assessment of confidence in the result?

6. TRANSITION.

The juncture between episodes is, in most cases, where managerial decisions (or their absence) will make or break a solution. Observe, however, that the presence or absence of assessment or other overt managerial behavior cannot necessarily be taken as either good or bad for a solution. In an expert's solution of a routine problem, for example, the only actions one sees may be reading and implementation. This explains, in part, the contorted and subjective nature of what follows.

TRANSITION questions:

a. Is there an assessment of the current solution state, and any attempt to salvage or store things that might be valuable in it?

b. What are the local and global effects on the solution of the presence or absence of assessment in part a? Was the action there appropriate or necessary?

c. Is there an assessment of the short and/or long term effects on the solution of the new direction, or does the subject simply "jump into" the new approach?

d. What are the local and global effects on the solution of the presence or absence of assessment in part c? Was the action there appropriate or necessary?

The Full Analysis of a Protocol

Appendix 3 presents the full protocol of two students working on the following problem:

> Consider the set of all triangles whose perimeter is a fixed number, P. Of these, which has the largest area? Justify your answer as best you can.

Student K is the same student that appeared in protocol 1. Student D (not the same as student D in protocol 2) was a freshmen with one semester of calculus behind him. This protocol was taken at the end of my problem-solving course, while protocols 1 and 2 were taken at the beginning.

The parsing of protocol 3 is given in Figure 3. The analysis given below follows that parsing.

Insert Figure 3 about here

Episode 1 (Reading, items 1, 2)

a. The conditions were noted, explicitly.

b. The goal state was noted, but somewhat carelessly (items 10, 11).

c. There were no assessments, simply a jump into exploration.

Transition 1 (Null)

a, b, c, d. There were no serious assessments of either current knowledge or of directions to come. These might have been costly, but were not--assessments did come in E_2.

Episode 2 (Exploration, items 3-17)

a. The explorations seemed vaguely goal-driven.

b. The actions seemed unfocused.

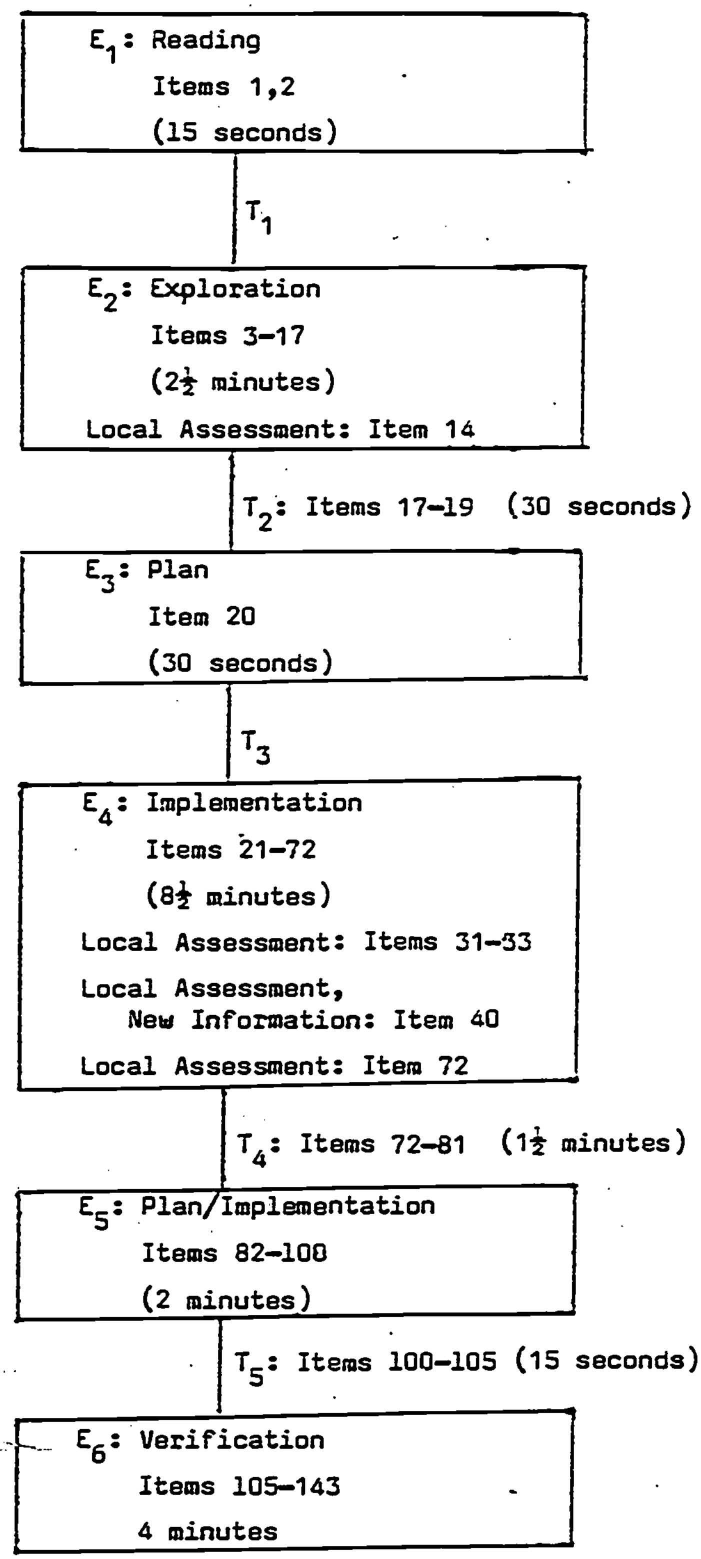


Figure 3

A Parsing of Protocol 3

c, d. There was monitoring, at items 14-17. This grounded the explorations, and led into Transition 2.

Transition 2 (Items 17-19)

a, b, c, d. Assessments were made both of what the students knew, and of the utility of the conjecture they made. The result was the establishment of a major direction: try to prove that the equilateral triangle has the desired property, and of a plan (episode 3). NOTE: If this seems inconsequential, contrast this behavior with the transition T_1 in protocol 1. The lack of assessment there, in virtually identical circumstances, sent the students on a 20 minute wild goose chase!

Episode 3 (Plan, item 20)

a. The plan is overt.

b. It is relevant and well structured. As to appropriateness and assessment, see the discussion of T_3.

Transition 3 (Null)

a, b. There was little of value preceding the plan in item 20; the questions are moot.

c. There was no assessment of the plan; there was immediate implementation.

d. The plan was relevant but only dealt with half of the problem: showing the largest isosceles was the equilateral. The "other half" is to show that the largest triangle must be isosceles, without which this part of the solution is worthless. . .a point realized somewhat in item 72, 8 minutes later. The result was a good deal of wasted effort. The entire solution was not sabotaged, however, because monitoring and feedback mechanisms caused the

<u>Episode 4</u> (Implementation, items 21-72)

a. Implementation followed the lines set out in episode 3, albeit in somewhat careless form. The conditions were somewhat muddled as the first differentiation was set up. The next two local assessments corrected for that (better late than never).

<u>Local Assessment</u> (Items 31-33)

1, 2, 3. The physically unrealistic answer caused a closer look at the conditions--but not yet a global reassessment (possibly not called for yet).

<u>Local Assessment, New Information</u> (Item 40)

1, 2, 3. The "new information" here was the realization that one of the problem conditions had been omitted from their implementation ("we don't set any conditions--we're leaving P out of that"). This sent them back to the original plan, without global assessment. The cost: squandered energy until item 72.

<u>Local/Global Assessment</u> (Item 72)

This closes E_4. See T_4.

<u>Transition 4</u> (Items 72-81)

a, b. The previous episode was abandoned, reasonably. The goal of that episode, "show it's the equilateral," remained. This, too, was reasonable.

c, d. They ease into Episode 5 in item 82. (It's difficult to say how reasonable this is. Had they chosen something that didn't work, it might have been considered meandering. But what they chose did work.)

<u>Episode 5</u> (Plan/Implementation, items 82-100)

c. They plunge ahead as usual.

d. The variational argument evolved in a semmingly natural way.

e. There was local assessment (item 95). That led to a rehearsal of the sub-argument (item 96), from which D apparently "saw" the rest of the solution. Further (item 100), D assesses the quality of the solution and his confidence in the result.

<u>Transition 5</u> (Items 100-105)

a, b, c, d. The sequel is most likely the result of a two-person dialectic. It appears that D was content with his solution (perhaps prematurely), although his clarity in explaining his argument in E_6 suggests he may have been justified.

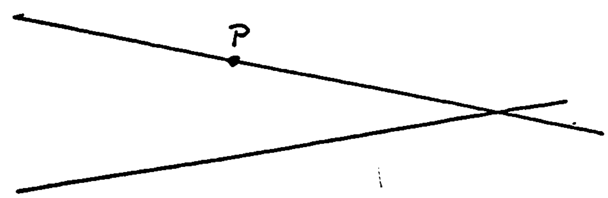
<u>Episode 6</u> (Verification, items 105-143)

This is not a verification episode in the usual sense. K's unwillingness to rest until he understands forced D into a full rehearsal of the argument and a detailed explanation, the result being that they are both content with the (correct) solution.

Some Empirical Results

Protocols 1 and 2 are relatively typical of the dozen protocols taken from pairs of students (six pairs, two problems for each pair) before a month-long intensive problem-solving course that focused on both tactics (heuristics) and strategies. The first problem was the one discussed in protocols 1 and 2, to find the largest triangle that can be inscribed in a circle. The second problem was a geometric construction:

> You are given two intersecting straight lines, and a point marked on one of them, as on the figure below. Show how to construct, using a straightedge and compass, a circle which is tangent to both lines and has the point P as its point of tangency to one of the lines.



Brief "snapshots" of a few representative pretest protocols are given below. These are too condensed to be useful for model building, but serve to demonstrate again the critical importance of managerial or strategic decision making. They also stand in (partial) contrast to the students' posttest behavior and (stark) contrast to some expert behavior. The diagrams that represent our episode analyses are here condensed into a sequential list of episode titles, with transitions deleted if there were none. Thus Figure 1 is rendered as (Reading/T_1/Exploration), etc.

E.T. & D.R., Problem 1. (Reading/T_1/Exploration)

After a brief mention of "max-min" problems, and a brief caveat ("But will it apply for all cases? I don't know if we can check it afterwards") in transition, they set off to calculate the area of the equilateral triangle. So much for the next fifteen minutes; in spite of some local assessments ("this isn't getting us anywhere") they continued those explorations. Result: all wasted effort.

E.T. & D.R., Problem 2 (Reading/Exploration)

In the initial explorations a series of sketches contains all the vital information they need to solve the problem, but they (without any attempt at review or assessment) overlook it. The solution attempt is undirected and rambling. Possibly because they feel the need to do something, they try their hand at an actual construction--already shown to be incorrect by their sketches--and are stymied when it doesn't work. Overall: lost opportunities, unfocused work, wasted effort.

Note: E.T. and D.R. are both bright; both had just completed the first semester calculus course with A's.

D.K. & B.M., Problem 2 (Read/Analyze/T_1/Explore/Analyze(Solve)/Verify)

Analysis is extended and coherent, but followed by a poor transition into an inappropriate construction that deflects the students off track for three and a half minutes. When this doesn't work they return to analysis and solve the problem. A detailed verification seals things up. Managerial decisions worked reasonably well here.

B.W. & S.H., Problem 2 (Reading/Exploration/T_1/Exploration)

A series of intuition-based conjectures led to a series of attempted constructions, the last of which happened to be correct--though neither student had any idea why, and they were content that it "looked right." This was a classic trial-and-error tape, and only because the trial space was small was there a chance that the right solution would be hit upon. There was one weak assessment (after a construction) that constituted T_1, but the result was simply a continuation of trial-and-error search.

Impetuous jumps into a particular direction were pretty much the norm in the pretests, and these first approaches were rarely curtailed. (This behavior was so frequent that it earned the name "proof by assumption," coined by my assistants.) Since there was little assessment and curtailment, little was ever salvaged from an incorrect first attempt, and a solution was often doomed to failure in the first few minutes of exploration.

Protocol 3, which has been discussed above, was taken after the problem-solving course. It is a representative, perhaps slightly better than average, sample of post-instruction performance. What makes this tape "better" than pre-test tapes is not that the students solved the problem, for their discovery of the variational argument that solves it may have been serendipitous. However, that they had the time to consider the approach was no accident: they had evaluated and curtailed other possible approaches as they worked on the problem. In general there was more evaluation and curtailment on the posttests than on the pretests, and less pursuit of "wild goose chases." In some cases this allowed for a solution, in some not; but at least their actions did not preclude the possibility. The following statistic summarizes the difference:

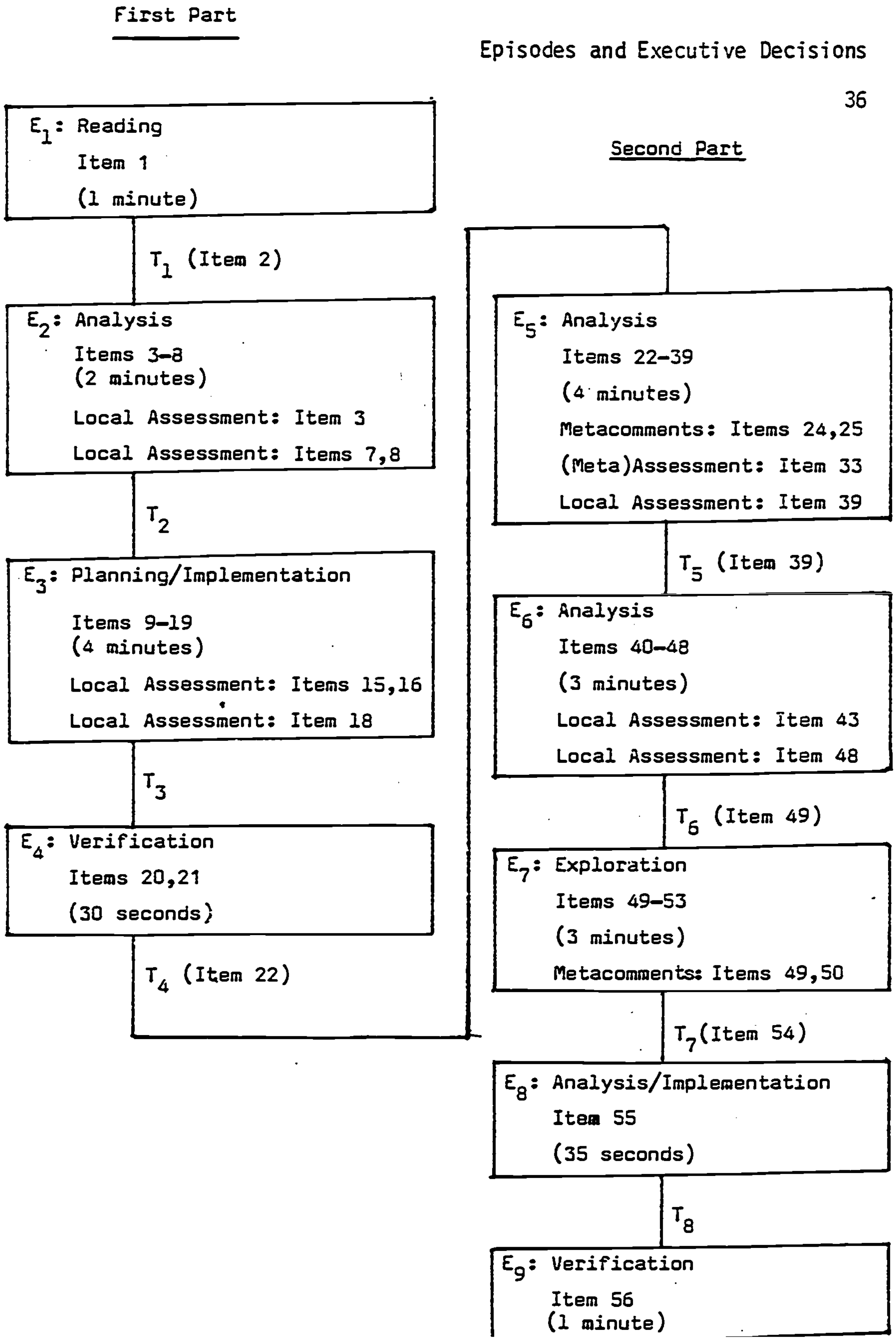
[illegible] of the [illegible] pretest protocols were of the type

Only two of the twelve posttest protocols were of that type. Not at all coincidentally, their performance improved on a variety of other measures as well (Schoenfeld, Note 7). However, the overall quality of the students' managerial monitoring, assessing, and decision making on the posttests was still quite poor. To indicate the contrast in managerial behaviors between experts and novices, we turn to the protocol of an expert working on a geometry problem. The expert, a number theorist, had a broad mathematical background but had not dealt with geometric problems for a number of years. It shows. By some standards, his solution is clumsy and inelegant. (In a department meeting it was held up for ridicule by the colleague who produced Protocol 5.) Precisely because the expert does run into problems, however, we have the opportunity to see the impact of his metacognitive, managerial skills.

The episode analysis of Protocol 4 is given in Figure 4. For (obvious) reasons of space, the full analysis will be condensed.

Insert Figure 4 about here

The critical point to observe in this protocol is that a monitor/assessor/manager is always close at hand during the solution. Rarely does more than a minute pass without some clear indication that the entire solution process is being watched and controlled, both at the local and global levels. The initial actions are an attempt to fully understand the given problem. By item 3 there is the awareness that some other information, or observation, will be necessary in order for a solution to be obtained. The actions in items 4 and 5 are goal-driven and, in item 6, yield the necessary information. This is utilized im-


First Part
E1: Reading
Item 1
(1 minute)
T1 (Item 2)
E2: Analysis
Items 3-8
(2 minutes)
Local Assessment: Item 3
Local Assessment: Items 7,8
T2
E3: Planning/Implementation
Items 9-19
(4 minutes)
Local Assessment: Items 15,16
Local Assessment: Item 18
T3
E4: Verification
Items 20,21
(30 seconds)
T4 (Item 22)
Second Part
E5: Analysis
Items 22-39
(4 minutes)
Metacomments: Items 24,25
(Meta)Assessment: Item 33
Local Assessment: Item 39
T5 (Item 39)
E6: Analysis
Items 40-48
(3 minutes)
Local Assessment: Item 43
Local Assessment: Item 48
T6 (Item 49)
E7: Exploration
Items 49-53
(3 minutes)
Metacomments: Items 49,50
T7(Item 54)
E8: Analysis/Implementation
Item 55
(35 seconds)
T8
E9: Verification
Item 56
(1 minute)

problem will be solved with one construction, which can be made. The plan is made in item 9. Implementation is interrupted twice with refinements (items 15 & 16; item 18) that again indicate that the subject is on guard for clarifications and simplifications at almost all times. The first part of the problem concludes with a quick but adequate rehearsal of the argument.

Like part 1, the second part of the solution begins with a qualitative analysis of the problem. In item 24, there is a comment that "this is going to be interesting" (i.e., difficult). Such a preliminary assessment of difficulty is, I believe, an indication of an important element of experts' metacognitive behavior. Experts seem to judge their work against a "template of expectations" when solving a problem. These expectations may be major factors in the experts' decisions to pursue or curtail various lines of exploration during the problem-solving process.

The solution of the second part continues, well structured, with a coherent attempt to narrow down the number of cases that must be considered. This is an implementation of "that kind of induction thought" from item 29. It appears to be a "forward" or "positive" derivation, verifying that all of the cases can be done. Yet the phrase "no contradiction" in item 33 reveals that the problem solver retains an open mind about whether the constructions could actually be implemented, and is still probing for trouble spots. The potential for a reversal, using argument by contradiction if he should come to believe one of the constructions impossible, is very close to the surface. This distanced overview, and the maintenance of a somewhat impartial perspective, are confirmed in item 49.

are planned ahead, but that the plans are assessed. Even the rather unusual excursion into quadratic extensions (item 53) is preceded by a comment about "knocking this off with a sledgehammer," and quickly curtailed.

In sum: this rather clumsy solution (see Protocol 5 in contrast), with its apparent meandering through the solution space, is in reality rather closely controlled. There is constant monitoring of the solution process, both at the tactical and strategic levels. Plans and their implementation are continually assessed, and acted upon in accordance with the assessments. Tactical, subject-matter knowledge plays a minor role here: metacognitive, "managerial" skills provide the key to success.

Discussion

This paper raises many more questions than it can answer. It was intended to. The extended discussions of protocols were designed to make one point absolutely clear: "metacognitive" or "managerial" skills are of paramount importance in human problem solving. As Brown observed (1978, p. 82), these types of decisions "are perhaps the crux of intelligent problem solving because the use of an appropriate piece of knowledge...at the right time and in the right place is the essence of intelligence." The inverse of this proposition should be given comparable stress: avoiding inappropriate strategies or tactics, at the wrong time or in the wrong place, is an equally strong component of intelligent problem solving.

To deal coherently with such executive decision making, one needs a framework for examining, modeling, and judging it. This kind of framework must, perforce, be substantially different from extant schemes like those used in mathematics education (Lucas, et al., 1979; Kantowski, Note 3), that focus on overt behaviors at a detailed level. As we saw in Protocol 1, the absence of an assessment may doom an entire solution to failure. Schemes that only seek overt behaviors cannot hope to adequately explain that protocol.

This kind of framework must also differ substantially from those used in Artificial Intelligence to simulate expert behavior in areas such as physics. Larkin, et al., (1980) characterize such work as depending on production systems to simulate the pattern recognition that "guide[s] the expert in a fraction of a second to relevant parts of the knowledge store...[and] guide[s] a problem's interpretation and solution (p. 1336)." While aspects of Protocol 4 such as the recognition of similar triangles (item 6) are compatible with this perspective, the whole of Protocol 4 stands in sharp opposition to it. At least half of the action in that protocol is metacognitive; it almost seems as if "manager" and "implementer" work in partnership to solve the problem. And it is precisely when the expert's problem-solving schemata (or "productions") do not work well that the managerial skills serve to constitute expertise.

The framework presented in this paper provides a mechanism for focusing directly on certain kinds of managerial decisions. Since a manager ought to be present at major turning points in a problem solution (if only to watch, in case action is necessary), the transition points between "episodes" are the logical place to look for the presence, or absence, of such decision making. Here we come to the first serious question: what, precisely, constitutes an "episode"? While there is reliability among coders in parsing these protocols at the macroscopic level, that begs the question: we need a rigorous formalism for characterizing such episodes. Unfortunately, I have not been able to adapt schemata for story understanding or for episodes in memory (see Bobrow and Collins, 1975) to deal with these kinds of macroscopic problem-solving episodes.

A formalism needs to be developed.

Questions regarding the characterization and evaluation of the monitoring, assessing, and decision making processes during problem solving are far more thorny. The role of the monitor was quite clear in Protocol 4; it assured that the solution stayed "on track." But how are these decisions made? It is clear from a variety of expert protocols that a priori expectations of problem or subtask difficulty serve as a basis for the decision to intervene. But the nature of the monitoring, the criteria for assessments, what the tolerances are, and how intervention is triggered all remain to be elaborated.

Similarly, assessment is not always desirable or appropriate: in a schema-driven solution, for example, one should simply implement the solution unless or until something untoward pops up. A simple-minded model that looked for assessment at each transition point between episodes (and other places) would miss the point entirely: assessment is only valuable some of the time, and we need to know when (and how).

In the long run, we need a detailed model of managerial monitoring, and assessment, and of the criteria used for assessment and decision making. This model will enable us to answer questions like those for the transition phase, "was the action or inaction appropriate or necessary?" In the meantime, these questions are not an evasion: they are an attempt to gather data so that the model can be constructed. A further refinement of these questions, and a much more detailed characterization of metacognitive acts in general, will be necessary. I hope that this paper provides a step in that direction.

Reference Notes

1. Ericsson, K. A. & Simon, H. A. Retrospective verbal reports as data. C.I.P. Working paper No. 388, 1978.
2. Ericsson, K. A. & Simon, H. A. Thinking-aloud protocols as data. C.I.P. Working paper No. 397, 1979.
3. Kantowski, M. G. The use of heuristics in problem solving: an exploratory study (Appendix D, Final Tech. Rep. NSF project SED 77-18543).
4. Kulm, G. Personal communication, "Process Code Dictionary," revised December 22, 1979.
5. Neches, R. Promoting self-discovery of improved strategies. C.I.P. Working paper No. 398, 1979.
6. Newell, A. On the analysis of human problem solving protocols. Paper presented at the International Symposium on Mathematical and Computational Methods in the Social Sciences, Rome, July 1966.
7. Schoenfeld, A. Measures of Problem Solving Performance and of Problem Solving Instruction. Manuscript submitted for publication, 1980.

References

Anzai, Y. & Simon, H. A. The theory of learning by doing. Psychological Review, 1979, 86(2), 124-140.

Bobrow, D. & Collins, A. Representation and Understanding. New York: Academic Press, 1975.

Brown, A. Knowing when, where, and how to remember: a problem of meta-cognition. In R. Glaser (Ed.), Advances in Instructional Psychology (Vol. 1). Hillsdale, New Jersey: Lawrence Earlbaum Associates, 1978.

Larkin, J. McDermott, J., Simon, D., Simon, H. Expert and novice performance in solving physics problems. Science, 20 June 1980, 208, 1335-1342.

Lucas, J. F., Branca, N., Goldberg, D., Kantowski, M. G., Kellogg, H., & Smith, J. P. A process-sequence coding system for behavioral analysis of mathematical problem solving. Columbus, Ohio: Ohio State University, 1979. (ERIC Mathematics Education Information Report Task variables in mathematical problem solving)

McDermott, J. & Forgy, C. Production system conflict resolution strategies. In D. A. Waterman & F. Hayes (Eds.), Pattern-directed inference systems. New York: Academic Press, 1978.

Neves, D. M. A computer program that learns algebraic procedures by examining examples and by working test problems in a textbook. In Proceedings of the second annual conference of the Canadian Society for Computational Studies of Intelligence, Toronto, 1978.

Polya, G. Mathematical discovery (Vol. 2). New York: Wiley, 1965.

Schoenfeld, A. H. Teaching problem solving in college mathematics: the elements of a theory and a report on the teaching of general mathematical problem-solving skills. In R. Lesh, D. Mierkiewicz & M. Kantowski (Eds.), Applied mathematical problem solving. Columbus, Ohio: Ohio State University, 1979 (ERIC Document).

Schoenfeld, A. H. Teaching problem-solving skills. American Mathematical Monthly, 1980, 87(10), 794-805.

Simon, H. Problem solving and education. In D. T. Tuma & F. Reif (Eds.), Problem solving and education: issues in teaching and research. New York: Lawrence Erlbaum Associates, 1980.

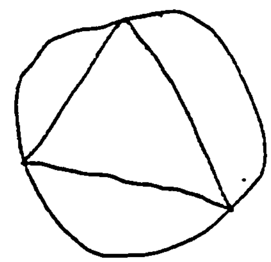
Stewart, P. Jacobellis v. Ohio. Decision of United States Supreme Court, 1964.

Appendix 1

Protocol 1

1. K: (Reads roblem) Three points are chosen on the circumference of a circle of radius R, and the triangle containing them is drawn. What choice of points results in the triangle with the largest possible area? Justify your answer as best as you can.

   You can't have an area larger than the circle. So, you can start by saying that the area is less than $1/2\pi R^2$.



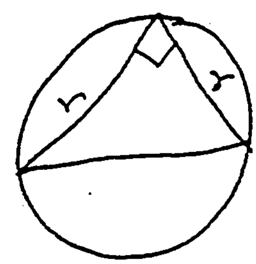
2. A: O.k. So we have sort of circle--3 points in front and R here and we have let's see--points--

3. K: We want the largest one--

4. K: We want the largest one--

5. A: Right, I think the largest triangle should probably be equilateral. O.k., and the area couldn't be larger than πR^2.

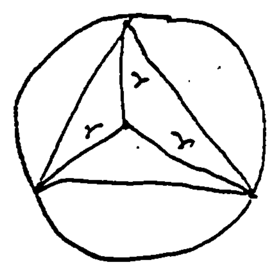
6. K: So we have to divide the circumference of the three equal arcs to get this length here. That's true. Right. So, 60-120 arc degrees--o.k.--so, let's see, say that it equals R over S--this radius doesn't help.



7. A: Do we have to justify your answer as best as you can? Justify why this triangle-----justify why you-----o.k. Right.

8. K: O.k. Let's somehow take a right triangle and see what we get. We'll get a right angle.

9. A: Center of circle of right triangle. Let's just see what a right triangle--is this point in the center? Yep, o.k. Yeah.



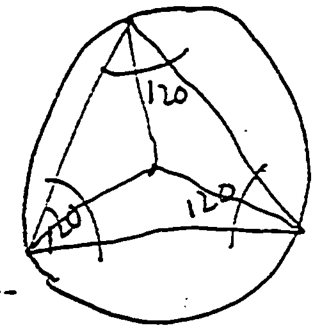
10. K: This must be the radius and we'll figure out that'll be like that, right?

11. A: So the area of this--

12. K: is R, is R--1/2 base times height, that's S and 2R, height is R so it is $1/2R^2$. It's off by a factor of 2.

13. A: O.k. But what we'll need is to say things like--o.k. Let's go back to the angle--probably we can do something with the angle.

14. K: Oh, I got it! Here, this is going to be 120--the angle of 120 up here--



15. A: Right! Yes, this is 120 and this is 120.

16. K: Right!

17. A: So--

18. K: We have to figure out--

19. A: Why do we choose 120--because it is the biggest area--we just give the between the biggest area--120.

20. K: Ummm. Well--the base and height will be equal at all times.

21. A: Base and height--right--

22. K: In other words--every right triangle will be the same.

23. A: Ah, ah--we have to try to use R, too.

24. K: Right.

25. A: O.k. (seems to reread problem)--justify your answer as best as you can. O.k. (pause)

26. A: So--there is the picture again, right? This is--both sides are equal--at this point--equal arc, equal angles--equal sides--this must be the center and this is the radius R--this is the radius R--

27. K: So we have divided a triangle with three equal parts and--

28. A: There used to be a problem--I don't know about something being square--the square being the biggest part of the area--do you remember anything about it?

29. K: No..I agree with you--the largest area...of something in a circle, maybe a rectangle, something like that...

30. A: Oh, well...so...

31. K: Since this is R--and this is going to be 120, wouldn't these two be R also?

32. A: Right.

33. K: This is 120.

34. A: Ah, ah.

35. K: Like a similar triangle--120 and 120 are the same angle--
so these two should be R.

36. A: O.k. Maybe they are.

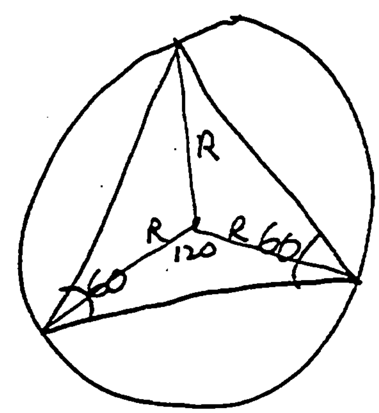
37. K: Why can't they be?

38. A: Mumbles------------

39. K: See, look--this is the angle of 120--right?

40. A: Right.

41. K: And this is an angle of 120. Right? This is like
similar triangles--



42. A: Wait a second--I think if you--this is true 120
but I don't think this one is-------

43. K: It is an equilateral triangle--that's--

44. A: No--it should be a 60.

45. K: That's right--it should be a 60.
Mumbles------that's 1/2 of it---that's right--
2R.

46. A: What are you trying to read from?

47. K: What if we could get one of these sides, we could
figure out the whole area.

48. A: Ah, ah.

49. K: Right?

50. A: Presume this to be 1/2 that side, we've got 1/2
base times height. We'll get the area--all we
have to show is the biggest one.

51. K: When we take the formula πR^2, minus 1/2 base
times height and then maximize that--then
take the derivative and set it equal to zero.
We can get that function--then we can get
this in the form of R.

52. A: O.k.

53. K: Then we can try this as the largest area.

54. A: Do you want to get this function, this as a function of R?

55. K: Yeah.

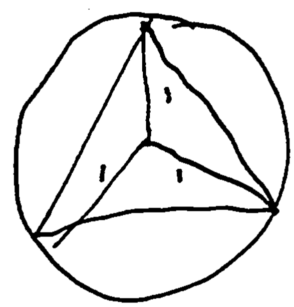
56. A: We can, I think. So you want this--right?

57. K: Well, it is kind of obvious that with B & H you are still going to have an R in it. So you can subtract it.

58. A: You have H in it. Well we have this one here. Mumbles--- (repeats the problem). Try this to be 2R.

59. K: No--it can't be. It has to be between R and 2R.

60. A: Yeah.



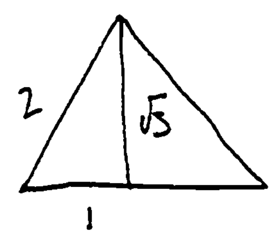
61. K: Helps us a lot! Set R equal to 1.

62. A: R = 1?

63. K: Right.

64. A: O.k.

65. K: That's one, that's one, that's one--it'll equal S over R. The area of the triangle is equal with R = 1, it's 2.



66. A: Well...height equals...

67. K: That's for the sides of the triangle--that's obvious -R = 1.

68. A: O.k.--divided into equal parts---(lots of mumbling)-- This from---well--you know--o.k. If you see we probably try to fix one point and choose the other two--o.k.--we are going to go from something that looks like this all the way down---

69. K: Right.

70. A: Right. O.k. and here the height is increasing where the base is decreasing.

71. K: Right. (Mumbles)

72. A: When we reach----o.k.

73. K: What is the area, side squared over 4 radical 2 for

an equilateral triangle? Is it like that?

74. A: You want the area for an equilateral triangle.

75. K: The area? I don't know. Something like side squared over radical 2, or something--

76. A: If you can probably show...at a certain point where we have the equilateral triangle the base and the...well...you know the product of the base since the base is decreasing and the height is increasing every time we move the line. If you can show a certain point, this product is the maximum--so we have the area is a maximum at that point. So this one is decreasing-----And at this point we have R, R, and R.

77. K: Ah, ah.

78. A: O.k. This is the base--is 2R--a right angle.

79. K: It wouldn't be $2R^2$.

80. A: Mumbles----One more--I mean--

81. K: O.k.

82. A: It should be R^2. But base times height--mumbles--and this one, say this is R + X.

83. K: The height equals R + X, so the base equals R-X.

84. A: Mumbles--those two things are equal to this--

85. K: Right.

86. A: All right.

87. K: I don't know.

88. A: We want this product of h as a maximum--as a maximum--and this one...I don't know.

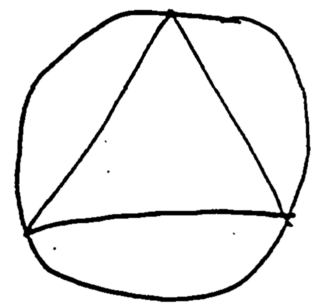
Appendix 2

Protocol 2

1. D: Reads the question.

2. B: Do we need calculus for this? So we can minimize, or rather maximize it.

3. D: My guess would be more like--mumbling--my basic hunch would be that it would be--

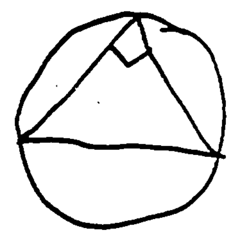


4. B: An equilateral--

5. D: 60, 60, 60.

6. B: Yeah.

7. D: So what choice of points has to be where on the triangle--these points are gonna be.

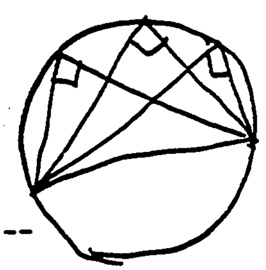


8. B: Try doing it with calculus--see if you can--just draw the circle--see what we'll do is figure out the right triangle--

9. D: Yeah, or why don't we find--or why don't we know the--some way to break this problem down into--like what would a triangle be for half the circle?

10. B: 60 degrees here?

11. D: Why don't we, why don't we say that--o.k.--why don't we find the largest triangle with base--one of the diameters, o.k.



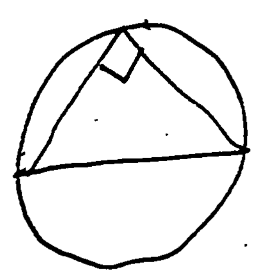
12. B: Base as one of the diameters?

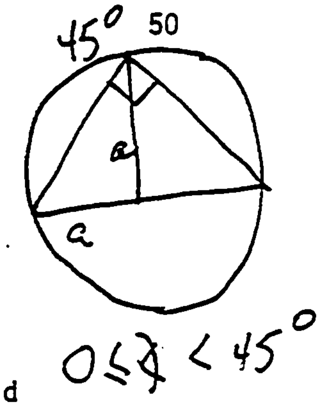
13. D: Yeah.

14. B: O.k. That would be just a family of right triangles--that go like this.

15. D: And they're all the same area?

16. B: No, no they're not all the same area--the biggest area would be in one like that. See if we could figure out--make it into sort of like a--if we could do it with calculus and I know there is a way. I just don't remember how to do it.





17. D: I have a feeling we wouldn't need the calculus. So this area then this is r and this would be--r^2--that would be the area of this--so then the distance here has got to be--45 degrees--

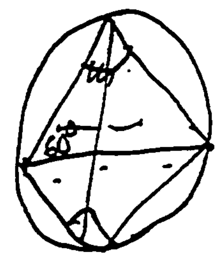
18. B: Right--that's got to be 45 degrees because they are the same. That's A--A over square root of 2--right?

19. D: Umma.

20. B: If that's radius--A--and this is A, too, so that would be A^2, that would be r^2, wouldn't it?

21. D: Right.

22. B: But I think this would be bigger.

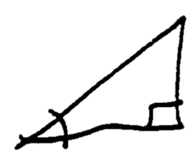


23. D: Oh, of course it would be bigger--I was just wondering if... (Pause)

24. D: Well we can't build a diamond--so we can't build a diamond that would go like that, obviously you want to make it perfectly symmetrical, but we can, if we maximize this area, and just flip it over, if we can assume that it is going to be symmetrical.

25. B: Yeah, it is symmetrical.

26. D: And if we can find the best area--



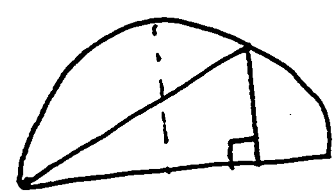
27. B: You mean the best--cut it in half in a semicircle.

28. D: Right. And if we can find the best area of--

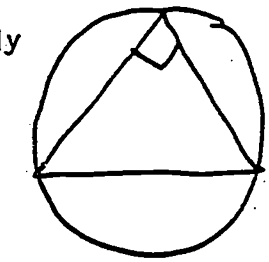
29. B: Any triangle that fits in a semicircle--well it wouldn't be a semi-

30. D: No it's a semicircle.

31. B: Largest triangle that fits in there?



32. D: Yeah, but it would have to be--if it is going to be symmetrical though, then you know this line has to be flat--it is going to have to form a right angle. So all we really have to do is form a right angle. So all we really have to do is find the largest area of a right triangle--inscribed in a semicircle.



33. B: Largest area of a right triangle. Yea, but obviously it is this one which is wrong.

34. D: No--No--

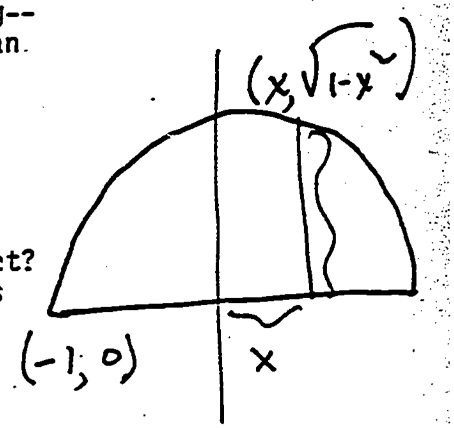
35. B: One like this.

36. D: Yeah with that angle, right.

37. B: O.k.--how we go about doing that? Hey, like we can--use the unit circle, right?

38. D: Umma.

39. B: So that means--this is $(1-x^2)$--this point right here--will be $(1-x^2)$, o.k. this squared--mumbling--I'll just put some points down to see if...pick an arbitrary--



40. D: Yeah, yeah, just to find this point--

41. B: All right, this is 1. Now I've got to find that point--o.k. What is the area of this--this is the distance right here times that distance, right? Product of those distances--area equals from this distance would be this, would be x value which would be x-1 or x+1? O.k., it's x+1, this distance right here times this distance right there which would be the y coordinate which is x^2. Want to take the derivative of that--to the x--mumbling.

$$A = \frac{1}{2}(x+1)\sqrt{1-x^2}$$

42. D: O.k.

43. B: Times (2-x). Did I have, oh, the [redacted] is crossed out so I just have an -x--or, that was [redacted]er $\sqrt{1-x^2}$, plus all this stuff. And set that equal to zero and you get that--oh, this is just one, isn't it--this is just one--so one of that, plus that equals zero, right?

$$\frac{dA}{dx} = (x+1)\left(\frac{1}{2}\right)\left(\frac{-2x}{\sqrt{1-x^2}}\right) + \sqrt{1-x^2}$$

44. D: I think we're getting a little lost here--I am not sure. Well, you go ahead with that--

45. B: Well, I'll just think about it, as it is just mechanical. There is a minus in here, isn't there? Mumbling--o.k. x equals $\sqrt{2}$ and what was this distance, we said? That was x--so that means it would be $\sqrt{2}$--plus 1--that's impossible.

$$\frac{1}{\sqrt{1-x^2}} = \frac{\sqrt{1-x^2}}{1}$$

46. D: Times R.

47. B: If x equals plus or minus the $\sqrt{2}$--

48. D: Umma--

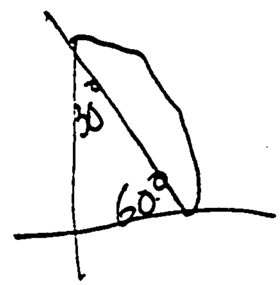
$\tilde{x} = -2$

$x = \pm\sqrt{2}$

49. B: This y thing would be 1 minus x^2, right?

50. D: This is just the distance--therefore, this right here has to be $\sqrt{2}$. Guess your calculations are all right.

51. B: Yeah, if I got x equals square root of 2--we've got a semicircle here, right? O.k.--and I have the points--right, it's a unit circle and I said that $x^2+y^2 = 1$, so $y = \sqrt{1-x^2}$. O.k.? And--(pause)--the x can't equal the square of the two because it would be out there. I know this has to be right but--



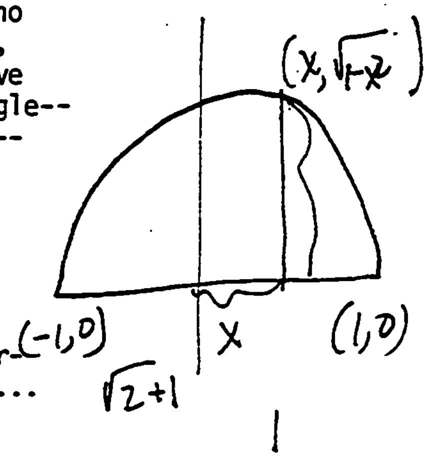
52. D: But all kinds of--let's see--well we know already, o.k. that the triangle is not 45, 45, because that would make it too small. O.k.?

53. B: Um--

54. D: So we know this angle is greater than zero and less than 90 degrees--

55. B: I just want to make sure I didn't--so this is x+1, x+1n...and cross multiply to set $1-x^2 = 1$ which means $x = \sqrt{2}$.

56. D: No, it has to be a 60, 60, 60--right triangle--no I am sorry not a right triangle--has to be a 60, 60, 60 triangle--because no matter where you move these vertices, it has to be a 60, 60, 60 triangle--because no matter where you move these vertices--



57. B: O.k.

58. D: --you are going to add area to this--like the--mumbling--you are going to add area to this.

59. B: All right, o.k. I understand, but I don't understand why it didn't work for this. I mean that... is there no solution for this equation?

60. D: I don't know--are you sure what you are looking for in that one?

61. B: Yeah. I marked off these and I just wanted to mark

62. D: O.k. What were you looking for? The length of this?

63. B: I was just looking for the maximum area of this--I said $A = (x+1)\sqrt{1-x^2}$. That's this height which is the square root of $(1-x)^2$. This is the unit circle. That's this distance right here--this minus the x value that I used--x value that is just x. O.k.--cause it is all in terms of x--x minus the x value here, which is x-1, which x+1--so area--ah shoot--I should have put 1/2 that is well,--mumbling--I'll get it. That should be 1/2 there, but I don't think that makes any difference--so that's all in terms of 1.

64. D: So--if--

65. B: Oh, wait a minute there's a difference--so one for two is 1/2 the first part--

66. D: So if you find the maximum area equal to--

67. B: It doesn't make any difference--it is just a factor of 1/2 here--because the area equals 1/2 that.

68. D: No--what's the next move?

69. B: See I get x--see I get a value of x with a plus or minus $\sqrt{2}$, right?

70. D: Umma.

71. B: If I plug x back into this I get $\sqrt{2+1}$, right? Then I plug x back into there and I get $(1-\sqrt{2})^2$ which is $\sqrt{1-1}$ which doesn't work.

72. D: Umma.

73. B: Which doesn't seem right. Plus r^2--mumbling--Let me just check my derivative over again. Now I know my mistake--hold it. I added this x--it's supposed to be times so we've still got a chance. So let me go from there. It is just a derivative mistake. Let me see it will be $(1-x^2)$--no it will be--$(-x+1)$. This might work--if it does--we solve that and cross out this minus 1. That means $x+1+x^2-1$, that makes x^2+x---cross this out--mumbling-- all right? It still doesn't work.

74. D: Well, let's leave the numbers for a while and

75. B: Yeah, you're probably right.

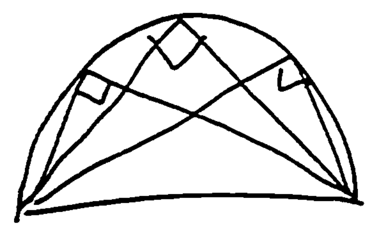
76. D: Well, we know that these two are some kind of symmetry.

77. B: Yeah.

78. D: I still say we should try--yeah--what we were doing before--just try to fix two of the points and let the third one wander around.

79. B: Yeah, we were going to fix them--yeah, I know what happens if you fix them on the diameter--then you have a family of right triangles.

80. D: Those the maximums.



81. B: Well, I don't see how--where are you going to fix the two points?

82. D: Well, you just fix them on any diameter. You find the largest triangle.

83. B: That would--obviously that would be the 45, 45 triangle if you fix them on the diameter. If you fix them on any chord.

84. D: Yeah, why though. Well, we know that if we put two of the points too close together--o.k.--o.k. --no matter where we put the third point--

85. B: Yeah.

86. D: --it's going to be too small. O.k. If we put them too far apart--o.k.--no matter where we put the third point. we are only using half a triangle.

87. B: O.k.

88. D: So it's got to be--o.k. So--two of the points, at least, well, matter of fact if you've got three points, each two of the points have to be between zero and 1/2 of the circle distance away from each other.

89. B: O.k.

90. D: See how I got that? O.k. so therefore each two of the points has to be like that--so

that? O.k. so we stick one point here--arbitrarily--so now the second point has to be somewhere o.k.--within--o.k. in other words, it can't be right here--it can't be right here--it can be anywhere else. We've got to place it so that the third point is going to be within half--

91. B: Half of what--I don't get you there.

92. D: O.k. Now wait a minute--let's see. You know when I said that--(pause). O.k. in other words the relationship between every pair of the three points....

At this point the interviewer (I) terminated the session and asked the students to sum up what they had done. B focused on the algebraic computations he had done in trying to differentiate $(1+x)\sqrt{1-x^2}$. The following dialogue ensued:

I: So what do you wind up doing, when you do that? You wind up finding the area of the largest right triangle that can be inscribed in a semicircle.

D: We determined that.

I. My question is: how does that relate to the original problem?

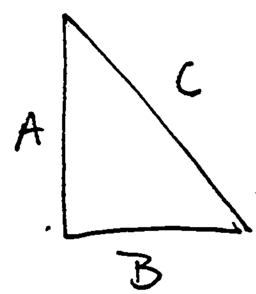
B: Well,....

Appendix 3

Protocol 3

1. K: (Reads problem.) Consider the set of all triangles whose perimeter is a fixed number, P. Of these, which has the largest area? Justify your assertion as best you can. All right now what do we do?

2. D: We got a triangle--well we know we label sides A, B and C.



3. K: Right. I'll make it a right triangle--all right--A,B, C and the relationship such as that 1/2AB = Area and A+B+C = P and $A^2 + B^2 = C^2$ and somehow you've got an area of one of these in the perimeter.

$\frac{1}{2}AB = A$

$A+B+C=P$

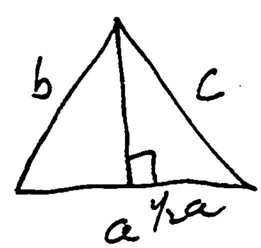
$A^2+B^2=C^2$

4. D: Yeah,except for somehow--I mean I don't really know--but I doubt that's the triangle of minimum area--well, o.k. we'll try it.

5. K: Largest area. Well, it is the only way we can figure out the area.

6. D: All right.

7. K: But for an isosceles we can do almost the same thing. This is 1/2(A). So that we know that the area is $(A/2)\sqrt{C^2-(A/2)^2}$. The perimeter = A + B + C and the height equals $\sqrt{C^2-(A/2)^2}$.



$b=c$

8. D: All right.

9. K: Now what do we do. We've got to figure out the largest area.

10. D: Isn't it the minimum?

11. K: The largest area.

$A=\frac{1}{2}(a)\sqrt{c^2-(\frac{1}{2}a)^2}$

12. D: So actually if we can get A--we have to get everything in terms of one variable and take the derivative, right? Basically?

$a+b+c=P$

13. K: Yeah,well--

$h=\sqrt{c^2-(\frac{a}{2})^2}$

14. D: Well, I still don't know if we should do--I

ever come to a problem like this--I mean we don't know--we have no idea as of yet with a given perimeter what's going to be that.

15. K: Right.

16. D: So, there--I mean--you can do that again but then what do you do?

17. K: Then we're stuck, right? Usually, you know, you could probably take a guess as to what kind of triangle it would be--like you could say it is a right triangle or an isosceles--I think it is an equilateral, but I don't know how to prove it.

18. D: Umma.

19. K: So we have to figure out some way to try to prove that.

20. D: All right, a good guess is that it is an equilateral, then why don't we try an isosceles and if we can find that these two sides have to be equal to form the maximum area, then we can find that--then we should be able to prove that side also has to be equal.

21. K: O.k. so B will be equal to C, so the perimeter P = A + 2B, or A + 2C = P.

22. D: All right.

23. K: Ummmm.

24. D: See what we've got.

25. K: Fix A as a constant then we can do this, solve that for C.

26. D: All right.

27. K: For a maximum area we've got 1/2, let's say A = 1, $\sqrt{C^2 - 1/4}$, right? Maximum area: $1/2(C^2-1/4)^{1/2} = 0$.

28. D: C^2 minus what?

$$\frac{1}{2}\sqrt{c^2 - \frac{1}{4}}$$

$$\frac{1}{2}\left(c^2 - \frac{1}{4}\right)^{1/2} = 0$$

$$\frac{1}{8}\left(c^2 - \frac{1}{4}\right)^{-1/2}(2c) = 0$$

$$2c = 0$$

$$c = 0$$

30. D: Ah, ah.

31. K: Mumbling--this is $1/4(C^2-1/4)^{-1/2}$. 2C, so we know that 2C has to = 0 and C = 0 and we are stuck!

32. D: We should have taken a derivative in it and everything, you think?

33. K: Yeah, that's the derivative of that. So does it help us? My calculus doesn't seem to work anymore.

34. D: The thing is--pause--you are letting C be the variable, holding A constant. So what was your formula--1/2 base times square root.

35. K: The base A times the square root times the height which is a right triangle to an isosceles which is --so it is $C^2-(A/2)^2$ which would give you this height.

36. D: $A^{2/4}$, no, $A^{2/2}$, no, $(A/2)^2$.

37. K: How about P = , ... no, C = P -A/2? Should we try that--

38. D: No, see part of the thing is, I think that for here we're just saying we have a triangle, an isosceles triangle, what is going to be the largest area? Largest area.

39. K: Largest area--set its derivative equal to 0.

40. D: All right. Well the largest area or the smallest area--I mean--if we are going to take a derivative--I mean--what's going to happen is you have a base and it's going to go down like that--I mean--we don't set any conditions--we're leaving P out of that.

41. K: Ah, ah.

42. D: That's absolutely what we have to stick in.

43. K: We've got C and a P-A over 2.

44. D: P -A over 2.

46. D: A + 2B = P--all right?

47. K: Shall we try that--mumbling. -A over 2--we've got to have a minus 1/4 PA--

$\frac{A}{2}\left(\left(\frac{P-a}{2}\right)^2-\frac{a^2}{4}\right)^{1/2}$

48. D: Well, then you can put A back in--then you can have everything in terms of A, right? Using this formula, we have the area and we have a--

$\frac{a}{2}\left(\frac{P^2-2a+a^2}{4}-\frac{a^2}{4}\right)^{1/2}$

49. K: All right--P--so that's $A/2\left(\frac{P^2-2A+A^2}{4}-\frac{A^2}{4}\right)^{1/2}$ and that's $A/2\left(\frac{P^2-2A}{4}\right)^{1/2}$...(mumbling and figuring)

$\frac{a}{2}\left(\frac{P^2-2a}{4}\right)^{-1/2}(2P-2)+$

50. D: Wait a minute--you just took the derivative of this right here?

51. K: This times the derivative of this plus this times the derivative of this.

$\left(\frac{P^2-2a}{4}\right)^{1/2}(1/2)=0$

52. D: Oh.

53. K: Mumbling and figuring...$A/4\left(\frac{P^2-2A}{4}\right)^{-1/2}(2P-2)+\left(\frac{P^2-2A}{4}\right)^{1/2}$ $1/2 = 0$...so $\frac{2AP-2A}{4}+\frac{P^2-2A}{8}=0$.

$\frac{2ap-2a}{4}+\frac{P^2-2a}{8}=0$

54. D: So can we get A in terms of P?

55. K: P^2--

$8P^2$

56. D: $8P^2 - 8P^2$ bring the P^2 on this side and multiply it by 8 and we'll have a qudratic in terms--no we won't--then we can just have A we can factor out in the equation--you see.

57. K: O.k. $P^2 =$

58. D: $-8P^2$--oh, are we going to bring everything else to the other side?

59. K: Yeah, 2A- +4A - 4AP x 8--No--

60. D: That's not right. Well, the 8 we can just multiply--

61. K: P^2 = all this.

62. D: Right

63. K: $P^2 - 4AP =$ --this isn't getting us anywhere.

64. D: $P^2 =$ factor out the A--then we can get A in terms of P.

65. K: $P^2 = 2A$--so you've got $A = \frac{P^2}{6+4P}$ --

66. D: So if we have an isosceles triangle and A has = to--

$$p^2 = 2a + 4a + 4ap$$

$$p^2 = 2a(3+2p)$$

$$\frac{p^2}{3+2p} = 2a$$

$$a = \frac{p^2}{6+4p}$$

67. K: be equal to that--

68. D: And if A has to be equal to that and B and C are equal--

69. K: So, B = --(whistles)

70. D: B = P- that.

71. K: $2B = P-A$ over 2.

72. D: No we aren't getting anything here--we're just getting--thing is that we assumed B to be equal to C so of course, I mean--that doesn't--we want to find out if B is going to be equal to C and we have a certain base--let's start all over, and forget about this. All right, another triangle. Certain altitude.

73. K: Well, let's try to assume that it is an equilateral.

74. D: All right.

75. K: Sides--mumbling--perimeter equals 3S, right?

76. D: Yeah, but wait a minute--that's still not going to really help us--what are we going to do--simply assume that it is an equilateral. We're just going to get that it is an equilateral, of course it is going to be an equilateral if we assume that.

77. K: True.

78. D: We want to prove that it is an equilateral if we think it is. If we want to do anything we can--

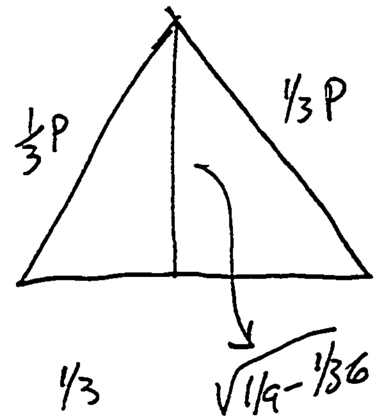
79. K: Yeah, how do you prove it?

80. D: Well, we can make up a perimeter--we don't need a perimeter P, do we? So,--

82. D: We want to maximize the area so that we can prove--o.k. we have the given base--we'll set our base equal to something.

83. K: Yeah, mumbling, P, or something--I don't know.

84. D: Then the other two sides have to add up to P.



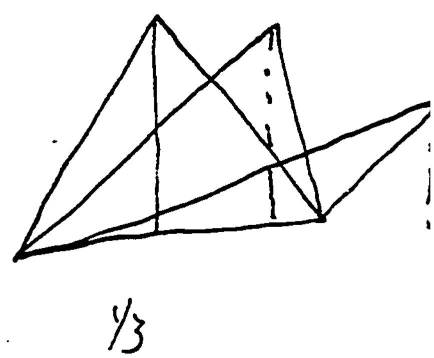
85. K: We--how about we say--let's start with an equilateral, just for the hell of it--see what happens. You get 1/3P, 1/3P and 1/3P. And this is 1/9 - 1/36 which is the height--

86. D: Now the thing we want to do is say--o.k. if we shorten this side at all and then what's going to happen to the height--if we leave this the same.

87. K: We can't shorten it.

88. D: And we shorten this side--sure we can--

89. K: Well--



90. D: We can have a--this equal to 1/3 and then a--this equal to--well you're going to have--I mean--

91. K: Aha.

92. D: This is going to get longer like that. Now we can see from this that all that is going to happen is that the base is going to get shorter so we know from that as far as leaving the base constant goes if we move--if we shorten this side then it is going to--somehow the point's going to go down in either direction.

93. K: Semicircle.

94. D: Right. That proves that we have to have an equilateral.

95. K: No, it proves an isosceles.

96. D: No, isosceles, I mean. All right from that if we set--we know that those two have to be equal so if we set this base equal to anything--it doesn't have to be 1/3P--we can also show that if this goes down--the area is going to get smaller, so

97. K: O.k., o.k.

98. D: In this case if it goes down to this side, we're going to have again a smaller angle here, shorter base here--and [noise].

99. K: So we get--so we know it is an equilateral--well prove it.

100. D: I don't know that's not a rigorous proof, but it is a proof--good enough for me.

101. K: Proves that an equilateral has the largest area.

102. D: Oh, we're talking about the largest area.

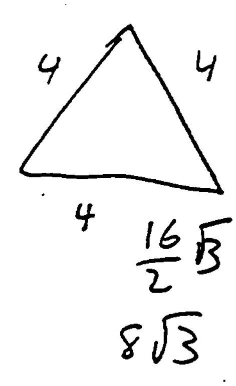
103. K: Yeah.

104. D: Oh, we just did.

105. K: We have to prove it has fixed number P--perimeter.

106. D: Well we already--we assumed that we have a fixed P, all right? I mean this is a proof as far as I.

107. K: Well, we've shown that an equilateral has the largest area. We haven't shown that if you have a certain set perimeter, let's say a right triangle, with a perimeter which is the same--we will not have a larger area.



108. D: No, but we have because we have shown with the set perimeter--o.k. we know that--

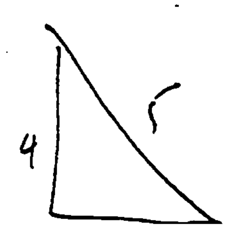
109. K: Well what if we have 3, 4, 5 with an equilateral being 4, 4, 4--

110. D: 3, 4, 5 is what? Mumbling.

111. K: 12. So this area will be 6 and this area will be side squared 16. --o.k. that will have the largest area.

112. D: What's that 1.7?

113. K: Yeah, 8 is still greater than 6 and that's greater than 1.



114. D: Oh, yeah, that's right. Yeah, but the thing is

side gets longer--say we use 4 as a base here, so then what's going to happen--well say we use 3 as a base, just so we won't have an equilateral when we are done--what's going to happen as 4 gets longer and 5 gets shorter--it's going to go upwards. The optimum area--the maximum area is going to be right there. Because you've got--

115. K: Right.

116. D: This angle and that height. If you make this angle any less--maybe let me draw a picture--

117. K: I can understand that--this will give us largest area, but how can we prove this bottom is one-quarter--1/3 the area of the perimeter?

118. D: Well, remember all the problems we've done where we say--o.k. let me just start from here once more--so that we have 3, 4, 5--is that what you have--because that's going to be 5. Wasn't a very good 3, 4, 5 anyway. So you start out with 3, 4, 5--all right, we pick the 3 has the base, right?

119. K: Aha.

120. D: All right, it's 5--mumbling--if we have 3 as the base--and this is a little bit off an isosceles, but if we draw an isosceles as 3 as the base--o.k. we've got a right angle--that's got to be the maximum--mumbling--(height?) because if it goes any--

121. K: Right.

122. D: Over this way, it is going to go down.

123. K: O.k.

124. D: All right, so remember the argument we've used--well if we--

125. K: Yeah, I can show that, but what you're not showing is--what you're not proving is that--

126. D: That it has to be an equilateral?

128. D: Right. I'm showing--first of all it has to be an isosceles. Right.

129. K: Right.

130. D: It has to be an isosceles--that means that we've got these three sides and those two are equal--right?

131. K: Umma.

132. D: Right--so now I pick this side as my base--I already picked--if that side is my base then the maximum area would have to have an isosceles--so I turn around--this side is my--

133. K: That I understand as proof, but you're not showing me that this is 1/3 the perimeter--mumbling.

134. D: If we have an isosceles triangle--if we have an equilateral triangle--then each side has to be 1/3 the perimeter--that's the whole thing about an equilateral triangle.

135. K: I know--o.k.

136. D: First we know it must be an isosceles, right?

137. K: Umma.

138. D: O.k.

139. K: I understand this.

140. D: If it is an isosceles, it must be an equilateral, right?

141. K: All right.

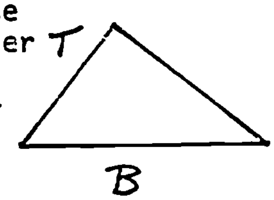
142. D: And if it must be an equilateral--all three sides must be equal and if the perimeter is P, all three sides must be 1/3P.

143. K: O.k. I've got it.

Appendix 4

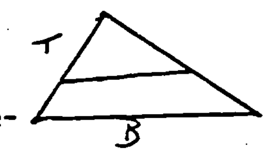
Protocol 4

1. (Reads problem) You are given a fixed triangle T with base B. Show that it is always possible to construct, with ruler and compass, a straight line parallel to B such that that line divides T into two parts of equal area. Can you similarly divide T into five parts of equal area?



2. Hmmm. I don't know exactly where to start.

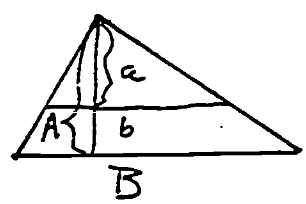
3. Well I know that the...there's a line in there somewhere. Let me see how I'm going to do it. It's just a fixed triangle. Got to be some information missing here. T with base B. Got to do a parallel line. Hmmm.



4. It said the line divides T into two parts of equal area. Hmmm. Well, I guess I have to get a handle on area measurement here. So, what I want to do...is to construct a line... such that I know the relationship of the base...of the little triangle to the big one.

5. Now let's see. Let's assume I just draw a parallel line that looks about right, and it will have base little b.

6. Now, those triangles are <u>similar</u>.



7. Yeah, all right then I have an altitude for the big triangle and an altitude for the little triangle so I have little a is to big A as little b is to big B. So what I want to have happen is 1/2 ba=1/2AB-1/2ba. Isn't that what I want?

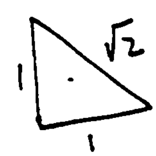
$$\frac{a}{A}=\frac{b}{B}$$

8. Right! In other words I want ab=1/2AB. Which is 1/4 of A times...mumbles(confused)...One over the square root of two times A times one over root two times B.

$$\frac{1}{2}b\cdot a=\frac{1}{2}A\cdot B-\frac{1}{2}ab$$

$$ab=\frac{1}{2}AB=\left(\frac{1}{\sqrt{2}}A\right)\left(\frac{1}{\sqrt{2}}B\right)$$

9. So if I can construct the square root of two, which I can! Then I should be able to draw this line...through a point which intersects an altitude dropped from the vertex. That's little $a=A/\sqrt{2}$, or $A=a\sqrt{2}$, either way.



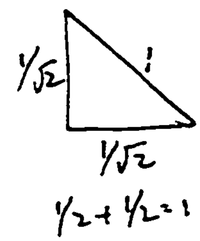
10. And I think I can do things like that because if I remember I take these 45^{o} angle things and I go $1,1,\sqrt{2}$.

$$\frac{1}{\sqrt{2}}A=a$$

11. And if I want to have a times root 2...then I do that...

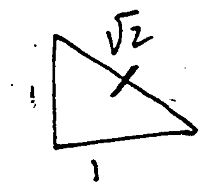
construct $1/\sqrt{2}$.

12. O.k. So I just got to remember how to make this construction. So I want to draw this line through this point and I want this animal to be...$1/\sqrt{2}$ times A. I know what A is, that's given. So all I got to do is figure out how to multiply $1/\sqrt{2}$ times it.



13. Let me think of it. Ah huh! Ah huh! Ah huh! $1/\sqrt{2}$...let me see here...ummm...that's 1/2 plus 1/2 is one...

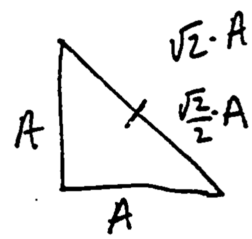
14. So of course if I have a hypotenuse of one...



15. Wait a minute: $1/\sqrt{2} \cdot \sqrt{2}/\sqrt{2} = \sqrt{2}/2$...that's dumb!

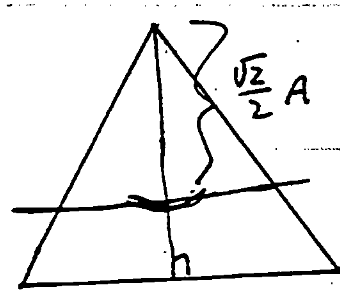
16. Yeah, so I construct $\sqrt{2}$ from a 45, 45, 90. O.k. so that's an easier way. Right?

17. I bisect it. That gives me root 2 over 2. I multiply it by A...now how did I used to do that?



18. Oh heavens! How did we used to multiply times A. That... the best way to do that is to construct A...A...then we get root 2 times A, and then we just bisect that and we get $A\sqrt{2}/2$. O.k.

19. That will be...what!...mmm...that will be the length...now I drop a perpendicular from here to here. O.k....and that will be...ta, ta...little a.

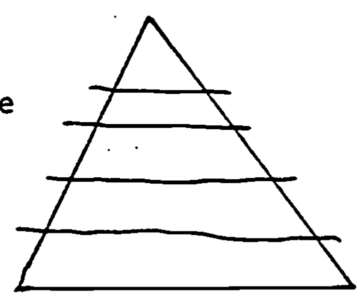


20. So that I will mark off little a as being $A\sqrt{2}/2$. O.k. and automatically when I draw a line through that point...I'd better get $\sqrt{2}/2$ times big B. O.k.

21. And when I multiply those guys together I get 2 over 4 times A times B. So I get half the area...what?...yeah...times 1/2...so I get exactly 1/2 the area in the top triangle so I better have half the area left in the bottom one. O.k.

22. O.k., now can I do it with 5 parts?

23. Assuming 4 lines.

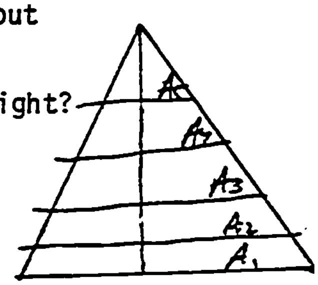


24. Now this is going to be interesting since these lines are going to have to be graduated...that...

25. I think, I think, that rather then get a whole lot of triangles here, I think the idea, the essential question is can I slice off...1/5 of the area...mmm...

26. Now wait a minute! This is interesting. Let's get a...how about four lines instead of...

27. I want these to be...all equal areas...right? A_1, A_2, A_3, A_4, A_5 right?

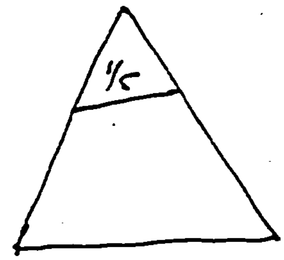


28. Sneak! I can...I can do it for a power of 2...that's easy because I can just do what I did at the beginning and keep slicing it in half all the time.

29. Now can I use that kind of induction thought.

30. I want that to be 2/5. And that to be 3/5.

31. So let's make a little simpler one here.



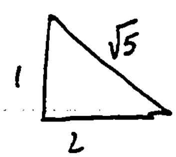
32. If you could do that then you can construct the square root of five. But I can construct the square root of 5 to one...square root of 5, right?

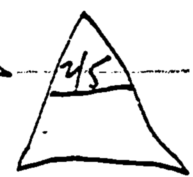
33. So I can construct...o.k. So that certainly isn't going to do it. No contradiction...

$a \cdot b = \frac{1}{5} A \cdot B = \frac{1}{\sqrt{5}} A \cdot \frac{1}{\sqrt{5}} B$

34. Now, I do want to see, therefore, what I have here.

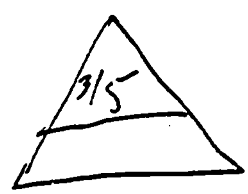
35. I'm essentially saying is it possible for me to construct it in such a way that that is 1, 2, 3, 4, 5, 1/5 the area...o.k.



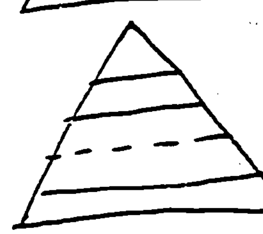


36. So little a times little b has got to equal 1/5 times A times B. So I can certainly chop the top piece off and have it be 1/5 of the area. Right? Right?

37. Now, from the first part of the problem...I know the ratio of the next base to draw...because it is going to be root 2 times this base. So I can certainly chop off the top two fifths.



38. Now, from the first part of the problem I know the ratio of the top...uh, o.k. now this is 2/5 here, so top 4/5... o.k....all right...so all I got to be able to do is chop off the top 3/5 and I'm done...



39. It would seem now that it seems more possible...let's see...

40. We want to make a base here such that little a times little b is equal to...the area of this thing is going to be 3/5... 3/5AB...in areas, right!...and that means little a times little b is $\sqrt{3}/\sqrt{5}$A times $\sqrt{3}/\sqrt{5}$B. O.k. then can I construct the square root of $\sqrt{3/5}$. If so then this can be done in one shot.

$ab = \frac{3}{5} A \cdot B = \left(\frac{\sqrt{3}}{\sqrt{5}} A\right)\left(\frac{\sqrt{3}}{\sqrt{5}} B\right)$

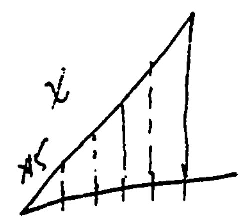
41. Well let's see. Can I construct $\sqrt{3/5}$. That's the question. $\sqrt{3}/\sqrt{5}\cdot\sqrt{5}/\sqrt{5} = \sqrt{15}/5$.

$\frac{\sqrt{3}}{\sqrt{5}}\cdot\frac{\sqrt{5}}{\sqrt{5}}=\frac{\sqrt{15}}{5}$

42. Root 15, root 15. Wait a minute! Root 15 over 5. Is the square root of 15 constructable? Root 15 is...

$\sqrt{16-1}$

43. It is the square root of 16-1. But I don't like that. It doesn't seem the way to go.



44. $16^2 - 1^2$ equals... (expletive deleted)

45. Somehow it rests on that.

46. (expletive) If I can do the square root of 15. Can I divide things and get this?

47. Yeah, there is a trick! What you do is you lay off 5 things. 1, 2, 3, 4, 5. And then you draw these parallel lines by dividing them into fifths. So I can divide things into fifths so that's not a problem.

48. So it's just constructing the square root of 15 then I can answer the whole problem.

49. I got to think of a better way to construct the square root of 15 then what I'm thinking of...or I got to think of a way to convince myself that I can't...umm...x^2-15.

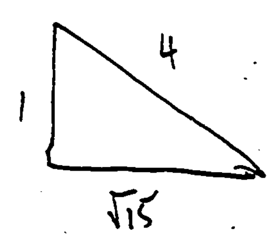
50. Trying to remember my algebra to knock this off with a sledge-hammer.

51. It's been so many years since I taught that course. It's 5 years. I can't remember it.

52. Wait a minute! Wait a minute!

53. I seem to have in my head somewhere a memory about quadratic extension.

54. Try it differently here. mmm...



55. So if I take a line of length one and a line of length... And I erect a perpendicular and swing a 16 (transcriber's note: for mathematical clarity he really means 4 instead of 16) here...then I'll get the square root of 15 here, won't I?

56. I'll have to, so that I can construct the square root of 15 times anything because I'll just multiply this by A and this by A and this gets multiplied by A divided by 5 using that trick. Which means that I should be able to construct this length and if I can construct this length then I can mark it off on here and I can draw this line and so I will answer the question as YES!!

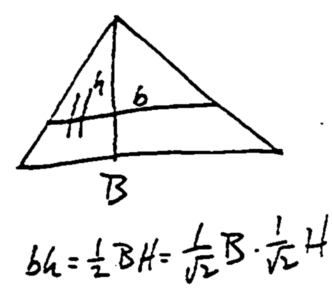
Appendix 5

Protocol 5

1. (Reads problem) Same as Protocol 4.

2. The first thought is that the two triangles for the first question will be similar.

3. And since we'll want the area to be one half. And area is related to the product of the altitude and the base we want the area of the smaller triangle to be one half.



4. And corresponding parts of similar triangles are proportional. We want the ratio of proportionality between the altitudes and the bases both to be $1\sqrt{2}$.

5. So I will draw a diagram...and I'm drawing that parallel and checking that algebra.

6. I hope you can hear the pencil moving because that's what's happening at this point.

7. And now I'm writing a bunch of letters on my diagram and multiplying them together...leaving the one half out, of course...and I want that to be one half of that.

8. So, that certainly seems like a reasonable solution. So all I have to be able to do is construct $\sqrt{2}$. And I can do that with a 45 right triangle, and then given a certain length, namely the altitude, to the base B, which I can find by dropping a perpendicular. I want to construct a length which is $1/\sqrt{2}$ times that, and I can do that with the ordinary construction for multiplication of numbers.

9. So, I can do the problem.

I: You can do all the constructions?

10. Yeah, I do them in the winter term. This line, this line, here's one, you want to multiply p times q, you draw these parallels and it's pq.

(The solution of part 2 is omitted)