DOCUMENT RESUME

ED 287 671 SE 048 625

AUTHOR Murray, Tom; And Others
TITLE Remediating Physics Misconceptions Using an Analogy-Based Computer Tutor. Draft.
INSTITUTION Massachusetts Univ., Amherst. Dept. of Physics and Astronomy.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Jul 87
GRANT NSF-MDR-840579
NOTE 50p.; Contains light and broken type.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College Science; *Computer Assisted Instruction; Concept Formation; Higher Education; Microcomputers; *Misconceptions; *Physics; *Science Instruction; Scientific Concepts; Student Evaluation; *Tutoring
IDENTIFIERS *Bridging Analogies; Science Education Research

ABSTRACT

Described is a computer tutor designed to help students gain a qualitative understanding of important physics concepts. The tutor simulates a teaching strategy called "bridging analogies" that previous research has demonstrated to be successful in one-on-one tutoring and written explanation studies. The strategy is designed to remedy misconceptions by appealing to existing correct intuitions, and extending these intuitions by encouraging analogical thinking. The strategy was motivating and successful for some learning situations. Analysis of the data indicates those situations in which the strategy worked well and those where alternatives are needed. Implications for improving the tutor using artificial intelligence technology to incorporate a representation of student beliefs and intelligent sequencing of example presentations are outlined. (Author/RH)

draft

Remediating Physics Misconceptions Using an Analogy-Based Computer Tutor[1]

Tom Murray[2]

Klaus Schultz[3]

David Brown

John Clement

Cognitive Processes Research Group
Department of Physics and Astronomy

University of Massachusetts
Amherst, MA

July, 1987

Abstract

We describe a computer tutor designed to help students gain a qualitative understanding of important physics concepts. The tutor simulates a teaching strategy called "bridging analogies" that previous research has demonstrated to be successful in one-on-one tutoring and written explanation studies. The strategy is designed to remedy misconceptions by appealing to existing correct intuitions, and extending these intuitions by encouraging analogical thinking. Students were videotaped while using the program and were encouraged to think aloud. The strategy was motivating and successful for some learning situations. Analysis of the data indicates those situations in which the strategy worked well, and those where alternatives are needed. We ou[illegible] implications for improving the tutor using artificial intelligence technology to incorporate a representation of student beliefs and intelligent sequencing of example presentations.

[1]Research reported in this paper was supported in part by the National Science Foundation Grant No. MDR-8470579

[2]Intelligent Tutoring Systems Group, Department of Computer and Information Science

[3]All four of the authors are also associated with the School of Education at the University of Massachusetts

1 Introduction

In this paper we report on a computer simulation of a human tutoring strategy which previous cognitive studies have shown to be effective in one-on-one human tutoring. The purpose of the research has been to evaluate the strengths and limitations of an automated version of the strategy, which we envision as ultimately being one of several strategies employable by an intelligent computer tutor. First we will place our work in the context of previous ITS research efforts.

Using artificial intelligence technology to build computer tutors can enhance learning in two ways. First, expert systems techniques combined with sophisticated knowledge representation methods facilitate the creation of learning environments which simulate complex processes or reify abstract systems and concepts. Second, AI technology facilitates the representation of human teaching expertise and pedagogical domain information. It can provide personalized feedback and assistance which is strategically relevant to the discourse context and the topic's pedagogical idiosyncrasies.

To date, most ITS research has focused on designing semantically rich learning environments or on the conceptual analysis of topic areas and tasks. Though these efforts are important, this paper addresses the paucity of research dealing with the act of tutoring itself. Though many intelligent tutoring systems incorporate expertise in teaching, diagnosis, and communication, few are *explicit* about the following: 1. what teaching strategies

are used; 2. assumptions made about the nature of learning and teaching which underpin these strategies; and 3. the precise rules or algorithms that instantiate these strategies. Sharing this information is conducive to the comparative analysis of tutoring system descriptions. It also encourages the proliferation of research that confirms or extends the findings of research designed to test or evaluate tutoring systems. Notable exceptions to the above concerns are the research efforts of Anderson et. al. (84), who investigates the learning process, Stevens & Collins (77), and Littman et. al. (86), who investigate the tutoring process, and Clancey (82) and Woolf (84), who investigate how teaching expertise can be explicitly represented in computer tutors.

As mentioned above, the computer tutor described here has been designed to simulate a human tutoring strategy. The strategy remedies misconceptions by appealing to existing correct intuitions, and extending these intuitions by encouraging analogical thinking. An overview of the strategy is as follows. The strategy is invoked when the student is diagnosed to be harboring a misconceptions, as evidenced by giving an incorrect answer to a target problem. The tutor then attempts to find an "anchor" situation. An anchor is a simple analogous situation for which the student has a correct intuition. For the domain we are investigating, an example target problem is the question of whether a table exerts an upward force on the book that rests on it. Many students exhibit the misconception that rigid objects can-

not exert forces. A typical anchoring situation is a book being held in a hand. Most students will agree that the hand exerts an upward force on the book. Typically, however, the student does not immediately see that the situations are truly analogous with respect to the questions being asked. Intermediate analogous situations, called "bridging analogies" are used to bridge the gap between correct intuitions and misconceptions. For example, a bridge between the book on the table situation and the book on the hand situation is a book on a spring. Bridging situations are presented one at a time to the student, chosen according to the student's previous responses, as described later. The student is asked to compare his answers about bridging situations with selected previous answers. These comparisons encourage the cognitive conflict that can motivate the student to see the problem situation as truly analogous to the anchor. The strategy has proven successful in remediating physics misconceptions in recently conducted human tutoring and written explanation studies.[4]

The computer "tutor" described here is not an intelligent program.[5] It is not intended to be used in a stand-alone fashion to teach. It is an encoding of a specific teaching strategy which research has shown to be effective in certain contexts. The purpose of the research is to evaluate the strengths and limitations of a computer version of the strategy. For this reason we

[4]The strategy has also been field tested in high school classes with encouraging results (Clement 87, and Camp et. al. 87).

[5]But see Section 5 for implications for an intelligent version of the strategy

have not embellished the program in ways which might make it teach better, if these embellishments would cloud the research questions. (However we did provide the user with features which made the program easy to interact with, such as graphics and an easy to use menu interface.) Serious investigation of teaching strategies before incorporating them into sophisticated computer programs is rare in ITS research. This study serves as a possible model for such investigations.

Human tutors seem to operate with a battery of tutoring strategies, many of which are relevant to specific contexts. They do not seem to use one general set of tutoring rules. We envision the bridging analogies strategy as one of many tutoring strategies that an intelligent computer tutor would have at its disposal, selecting each (or combinations of them) according to context applicability rules. We see these strategies as "strong method" solutions to the computer tutoring problem because their appropriateness is context dependent, relying on information about the student, the discourse state, and the topic characteristics. Tutoring guidelines or strategies mentioned in previous research are usually "weak methods" in that they are more generally applicable, but not as powerful as context sensitive strategies. Examples of weak method strategies are: "provide concrete examples of, as well as descriptions of, concepts" (Burton & Brown 82); "provide immediate feedback for errors" (Anderson et. al. 84); and "let the student learn from mistakes" (Brown, Burton, & deKleer, 82).

Ideally, the effectiveness of each of the strategies employable by a computer tutor should be researched. Our view is that it is desirable to test such strategies independently, first off line, and then in a prototype program. In this way, a great deal is known about these strategies before a large effort is expended in building an Intelligent Tutoring System based on them. Of course, further research is then needed to determine how to integrate several tutoring strategies into a single system.

Key research questions

At the onset of the research there were several key questions:

- Can the strategy be reasonably simulated in a computer environment?
- Can the strategy effectively remediate misconceptions?
- Under what circumstances (i.e. what types of students, types of domains, discourse situations, etc.) does the strategy work?
- What aspects of the strategy were most and least effective?

Longer range questions that we are addressing are:

- What other tutoring strategies might be effective where bridging analogies fails?
- Is it possible to determine dynamically during a computer tutoring session when to use the bridging analogies strategy and when to switch

to another strategy?

The bridging analogies strategy presents the student with situations that bridge the conceptual gap between simple situations (those for which most students have correct intuitions) and difficult ones (those for which most students exhibit misconceptions), and encourages analogical thinking in an attempt to transfer the students' correct intuitions to a broader context. The strategy is designed to be applicable to topic areas with a qualitative, conceptual focus, where deep seated misconceptions exist, and where correct intuitions (anchors) also exist. Basic Newtonian mechanics is one such domain. Real world experiences result in a wealth of intuitions about how the physical world works. Some of these are in accordance with scientific laws, and others in conflict with them.

2 Background—Misconceptions in Science

Research studies in recent years have investigated students' pre-conceptions about important mathematics and science concepts prior to formal instruction. A surprising number and variety of misconceptions (incorrect preconceptions) have been uncovered in domains such as electricity, optics, biology, statistics, and classical mechanics. Often these misconceptions are deep-seated and counterproductive to problem solving in the domain. Such incorrect beliefs are not limited to non-science students. Serious misconceptions are common among students taking calculus-based physics courses and

students who are proficient at quantitative problem solving using formulas. One implication for instruction is that the existence of these debilitating beliefs must be dealt with explicitly. They are not simply "erased" when conflicting and correct knowledge is conveyed to the student. Many of these misconceptions are commonly held, so that attempts to define and remedy the most common of them will benefit a large percentage of students. For instance, in a survey of 112 high school chemistry and biology students who had not taken physics, Clement (87) found that 76% of the students did not believe that a table exerts an upward force on (or pushes up on) a book resting on it (83% of these indicated a high confidence in their incorrect answer), a belief that is contrary to Newton's action-reaction law.

Unfortunately, these misconceptions have proven to be quite resistant to remediation, at least via traditional instructional methods. Several studies support this finding, the most comprehensive by Halloun and Hestenes (1985). They administered tests of basic concepts in Newtonian mechanics (motion and its causes) to over 1000 college level introductory physics students over a three year period. The primary diagnostic instrument consisted of multiple choice questions about very basic concepts, such as: "if a ball is shot out of a semi-circular tube, what is the shape of its path after leaving the tube?" (the student was to choose between pictures of five path shapes). Average scores before instruction were "very low," (about 45%) even for university physics courses where 80% are engineering majors,

and 80% have taken calculus. Most importantly, the *post-instruction* scores showed little improvement (11 to 15%). This scant gain was remarkably independent of such variables as instructor, math ability, and final course grade.

The literature cites several possible sources for students' misconceptions. White & Frederiksen (86) note that misconceptions can result when the cognitive jumps required of a student while learning a new domain are too large or too under-specified. Learners can then fill in the gaps in unpredictable and uncontrollable ways. Such beliefs, acquired during instruction, might be called "mis-comprehensions." Similarly, diSessa (1985) discusses how misconceptions can result when learners try to accommodate to new information by incorrectly synthesizing existing fragmented and inconsistent pieces of knowledge. VanLehn (83) has constructed a detailed theory of how procedural knowledge is modified ("repaired") in an ad-hoc way when an impasse is encountered during problem solving. A similar mechanism may account for some (non-procedural) misconceptions. Claxton (85) points out that preconceptions can be founded on physical experiences ("gut science") or on social experiences ("lay science"), and that these concepts can conflict with the formal and abstract concepts of "school science." Students are often perfectly content with or unaware of their conflicting beliefs if the contexts in which these beliefs are used do not overlap, such as in everyday life vs. taking exams.

Many misconceptions result from repeated attempts at comprehending real world phenomena. Some of these beliefs have been used again and again to successfully cope with or explain real world events. Therefore it is not surprising that some misconceptions are deep seated, and that some are quite common. Some domains, such as physics (McCloskey 83 and McDermott 84) and statistics (Tversky & Kahneman 74), are particularly susceptible to the existence of misconceptions prior to any formal instruction. Interestingly, Clement (1982) has documented remarkable similarities between students' misconceptions and pre-Newtonian theories such as those of Galileo. In such domains, careful sequencing and explanation of new information is probably not sufficient, and existing misconceptions must be directly addressed with innovative instructional approaches.

3 Description of the Bridging Analogies Tutor

Clement & Brown (84) have developed a teaching strategy, called "bridging analogies," which utilizes correct intuitions (which they call "conceptual anchors") and appeals to students' analogical reasoning to help them extend their correct intuitions to target situations for which they have misconceptions. After the student gives an incorrect analysis of a situation that indicates the existence of a common misconception (such as that rigid objects don't exert forces), she is presented with an analogous anchor situation (such as a book in a hand), which we will assume is indeed an anchor to

the student.[6] The student has given contradictory answers for the two situations. Apparently she does not see them as analogous in terms of their relevant properties. The bridging analogies strategy attempts to bring the student to an understanding of this analogical relationship by presenting a sequence of intermediate analogies (called "bridging analogies"). At some point (or points) the student should be faced with considering two situations for which she has given contradictory answers yet which she realizes are analogous. The cognitive conflict which results should motivate her to change her mind about the misconceived situation.

The strategy employs Socratic type questioning to engage students in thought experiments. Decisions of whether to bridge are based on student responses, so that errors, rather than a pre-defined hierarchy of prerequisite information, provide the catalyst for instruction. The strategy delays revealing new information about the subject matter and giving feedback about the correctness of student responses as long as possible. This encourages the construction of new ideas based on information which the student already has. This pedagogy, if taken to the extreme, is of course limited. The discourse can stray so far from the goal that consolidation becomes dif-

[6]In a classroom situation anchors must be chosen which are expected to be intuitive to the majority of the class. In one-on-one tutoring, an anchor can be found with a series of questions. In either case, previous research must establish which situations are useful anchors for a given domain.

ficult, and some students become frustrated without a certain frequency of informative feedback. But withholding feedback can have many advantages as well. For most students, who are used to memorizing correct answers, immediate feedback can inhibit deeper thought about conceptual issues. In instructional dialogues where feedback is withheld there can be more emphasis on the quality of the ideas themselves than on getting correct answers to questions. Students are forced to evaluate their own knowledge and learning processes. This instructional approach is based on the assumption that students must be actively engaged in constructing a new conception in order to replace their misconceptions. This research will help determine the effectiveness and limitations of this pedagogical style.

For computer implementation, the strategy was formulated as a doubly recursive procedure which traverses a pre-defined network of example situations according to student responses to questions.[7] The computer version of the strategy cannot respond to creative student insights, but it can tailor its presentation of examples personally for each student. We have used the program for different domains, but below we describe its implementation in teaching about the existence of opposing contact forces in static situations. This is the domain used for the majority of the research, and for which the most is known.

[7]In this section we give an operational definition of the procedure. See Appendix A for a more precise description of the code implementing the strategy

Figure 3: Diagram of the bridging analogies strategy

bridge (the book on the spring) is (apparently) remedied (see Figure 3), and the anchor-to- bridge analogy is established.

Once the bridge situation is understood and the anchor-to-bridge analogy established, we wish to establish the bridge-to-problem analogy, thus constructing a path of analogies or conceptual links connecting the anchor with the problem. The strategy compares the bridge (the book on the spring in this case) with the problem (the book on the table) with the hope that the student will change her mind about the problem. If the student does so we are done (or done with the current recursive call of the algorithm). If the student does not opt to change her mind about the problem, we are back in our original condition, having a misconceived situation (the book on the table) and a correctly conceived one (the book on the spring)—which this time is the bridging situation. The algorithm is called recursively to establish the bridge-to-problem analogy, beginning by finding a new bridge (in this case the book on a flexible board) and so on.

One perspective on the strategy is that is keeps splitting the conceptual difference between the simplest misunderstood situation and the most difficult understood situation. Eventually it provides the student with an intuitively valid analogical path from the original anchor to the original problem, such that the gaps between neighboring situations are close enough for

the student to perceive the analogy. Splitting the difference at each step is more efficient than walking the student through every smallest step of the entire set of example situations linking the problem with the anchor. The dark line in Figure 1 shows one student's traversal through the network. Since the strategy is "middle out", the path shows the order that situations were understood, rather than the order in which they were introduced.

Hints for example situations

Available for every example situation are:

- a description of the situation
- optionally, a "detailed description" (as described below)
- a question with multiple choice answers
- a hint or explanation of the correct answer
- and, optionally, a pointer to a graphics file for the picture.

The system does not distinguish between hints and explanations. The ones used in the static forces network varied along a continuum from weak hints to strong hints to correct answer descriptions. There are two circumstances where the hint or correct explanation is given to the student. One is when the student, after a comparison, opts to change her mind about an anchor and changes it from right to wrong. At this point we have no

more graceful method for dealing with these "regressions". They occur very rarely in our tests so far.

The second condition warranting a hint or explanation (i.e. the student gets immediate feedback after a wrong answer) is when there is no bridge in the network between the current problem and anchor. The hint is given, often suggesting that the two situations are analogous, and the student is asked the problem question again.

Confidence checking

In order to simplify the description of the bridging analogies strategy, we have thus far neglected an important feature: confidence checking.

A severe drawback to computer tutoring, as compared with human tutoring, is its inability to ascertain the types of affective information about the student that are evident to human tutors from facial expressions, vocal inflections, etc. Information about students' certainty level, confusion, boredom, excitement, etc., are crucial to effective human tutors.

Human tutoring studies at the University of Massachusetts have used "how sure are you" or, alternately, "how much sense does the answer make" scales to gather additional information about the strengths of anchors and the effectiveness of the bridging strategy in changing students' beliefs. Borrowing on this idea, the computer tutor asks the student to rate her confidence in her answer on a five point scale after each question (as shown in

Figure 2). The program combines this information with information about the correctness of the answers in deciding whether to ask for comparisons, and whether to bridge (see the section on data structures and algorithm for details).

As mentioned above, the strategy postpones any type of feedback concerning the correctness of a student's answer as long as possible.[12] The algorithm is sufficiently complex so that the presentation of situations and comparison questions appear random to the user. Depending on the program's internal state in the recursive procedure calls, a new situation or comparison could be given after either a right or wrong student response. Therefore, the student usually does not know why she is presented with a new example or asked to compare situations. This tends to keep the student focused and thinking about the current physics question, rather than second-guessing the tutor's intentions. The student tends to answer according to what she believes, rather than what she thinks the tutor wants her to believe (as can happen when students have an attitude of getting the most "points" rather than an interest in thinking and learning).

Since confidence information is important, the comparison allows students to change their confidence about a previous answer as well as the

[12]This strategy is in some aspects extreme. The reader should keep in mind that we are conducting research on specific strategies in isolation, to assess their effectiveness in limited instructional contexts.

answer itself. We have noted in some computer tutoring sessions that asking students to commit to a confidence rating causes them to put additional cognitive effort into visualizing and analyzing the example situations.

Topics

Example situations are grouped into "topics" corresponding to the physics concepts to be taught. The topic's grain size is relatively small, and each topic has a network of examples. The topic for the network shown in this paper is "existence of contact forces in the vertical direction between two stationary objects." For each topic area the goal of the tutor is to help the student understand the qualitative physics of a key "target situation" which exemplifies an expert application of the concept. If the student answers incorrectly about a target situation, then the bridging analogies strategy is invoked. The topics are given in sequence, and if there is any overlap between example situations in the topic areas, the program takes into account previous answers about the situations. In this paper we describe results for only the topic network shown in Figure 1.

Designing the example situation network

The network design can be described in two ways: in terms of example difficulty, or in terms of example features. We will discuss each below. It should be kept in mind, though, that the final version of the network we

used was influenced by empirical results from human tutoring studies.

Traversal along the extremities of the net will touch each example situation, yielding a sequence that can be interpreted as, starting with the original anchor, progressing from easier to harder problems. Each bridge is intended to split the conceptual distance between two other situations. The bridging analogies strategy is seen in this light as efficiently searching the space of examples using a divide and conquer approach. The sequence of examples between the current problem and the current anchor is divided at the bridge specified by the network. The network thus encodes the pedagogical knowledge of estimated conceptual distance (as opposed to distance in a linear sequence of examples, where the third would be half way between the first and fifth).

A second perspective describing the network involves feature dimensions of the examples, and is an equally valid description for the book on the table network we have been using. Research has uncovered several reasons for students' beliefs that the table can not push up on the book, but a hand *can* push up on a book. Each reason involves an attribute which distinguishes the two examples. Reasons include: a person is *doing* something, but the table is just in the way; a person's arm can move to adjust to the book's weight, but the table cannot; an arm's purpose is to exert forces, but a table's is not. Thus the student sees one or several features of objects as being relevant to their ability to exert forces, all of which are irrelevant from a physicist's

standpoint. The network factors out these possible features incrementally by providing example situations with different sets of relevant and irrelevant features. For example, the spring can move and is designed to exert a force, but can not decide on its own to volition to do so.[13]

Other features of the tutor

The tutor is written in PASCAL and runs on an IBM PC. As was mentioned above, the domain specific information and inter-example links can be easily modified, or new domains created from scratch, by straightforward word processing of text files. However, it is not easy to design *effective* networks of anchors and bridges. This requires research iterations with human subjects.

Figure 2 shows examples of the graphic pictures that accompany the descriptions of the example situations. Created using the PC-Paint (trademark of Mouse Systems) program, their inclusion is optional.

All student input is through menu selections. The use of menu-driven options, and of multiple choice answers, allows us to bypass difficult problems of language recognition involved in interpreting student input. Multiple choice answers can include reasons for the answers and "distractor" items which catch common wrong answers and misconceptions. Usually, reasons for a student's answer are just as important as the correctness of the answer.

Example situations can have two levels of description associated with

[13]See the section on Implications for a more formal approach.

them. After being presented with a description of a new situation, the student is asked if she wants a more detailed description. The detailed description attempts to address possible areas of confusion which might arise.

4 Formative Evaluation of the Tutor

The purposes of this phase of evaluation are, as mentioned, to determine what aspects of the simulated strategy are effective, and in what instructional contexts. We are also interested in information leading to alternative computer tutoring strategies to be used where bridging is not effective.

The first working version of the tutor was quite primitive. We were not sure whether a machine simulation of the human strategy would be so different as to make previous results concerning the usefulness of the strategy irrelevant. A pilot study was conducted with an interviewer showing the subjects picture cards corresponding to the text on the computer screen. Evaluation of the videotaped student interactions confirmed that an automated version had promise, and we began a second design cycle, which is the focus of this paper. The pilot study provided information leading to several modifications to the program. The user interface was revised, the format of the comparison question was changed, graphics capabilities were added, and several nodes were added to the book on the table example network.

To locate subjects suitable for an interview study of the revised tutor, a

pretest was administered to students in several university classes. A total of 180 students were given a pretest with four example situations, questions, and confidence ratings. None of the questions were the same as the ones in the tutor's network. One situation was similar to the original problem, involving contact forces on a medium sized object resting on an inflexible object. Another was similar to the initial anchor, where one object was held in a hand. 84 percent of the students answered the suspected anchor correctly, and 68 percent answered the suspected problem incorrectly. 63 percent got both the problem wrong and the anchor right. These figures are in rough agreement with the results of previous studies.

Of the students who took the pretest, 53 percent volunteered to be subjects for a one hour video taped interview study, for which they would be remunerated 5 dollars if chosen.[14]

Twenty five students were selected for interviews. The selection of students was biased toward those for whom we thought the strategy, and the network used, were most appropriate. Those who answered the problem incorrectly with high confidence and the anchor correctly with high confidence were given preference. Three of the sessions involved pairs of students using the tutor. There were technical problems (including power outages), and several students unexpectedly answered the original problem correctly with

[14]The distribution of correct answers for the two situations did not differ significantly between those who volunteered and those who didn't.

high confidence right from the start—in which case the bridging analogies strategy is not invoked at all (or needed). In the end we had a total of 15 sessions with useful data from the book on the table network.

Summary of the results

Several types of data were recorded (see Appendix B for a description of how the data were analyzed). An interviewer, who was present to elicit thought verbalization and answer student questions about using the program, [15] took notes during the sessions. We have completed a first pass analysis of the video taped recordings. See Schultz et. al. (87) for a protocol analysis report. The program stores a trace of each session, and this information has been summarized and analyzed. Following are some observations:

- At the end of the sessions all but one of the 15 subjects had a high confidence in the correct answer to the original problem situation, and most of their verbal reasons indicated that the correct answer made sense to them at an intuitive level. Nine of them started the session with a high confidence in the wrong answer to the original problem. (The rest started with a low confidence in the wrong answer to the original problem.)

[15] The current version of the program is not robust enough to be used without supervision.

- Most of the subjects needed to pass through a node which described all matter as being composed of molecules connected by springy bonds. Since we designed the network, we have come to categorize this type of example as a "causal model" (Brown & Clement 87), rather than a bridging analogy. It is not clear whether their misconception concerning the book on the table could have been remedied for these subjects without the use of a causal model.

- The analogical comparisons alone successfully instigated a change in belief at some point in the network for 50 percent of the subjects. This suggests that bridging and analogical thinking are effective instructional strategies for many students at least locally (i.e. if the analogy is not too distant), but perhaps not when used exclusively.

- Nine of the subjects needed at least one explanation or hint at some point in the session. This could indicate that no bridge was available where the student needed one, or that the bridging analogies strategy is not powerful enough to be used exclusively in computer tutoring in these types of domains.

- There were only two instances of "regression," where a student, after a comparison question, attempted to change her mind about an anchor from right to wrong. Also, there was only one case where a student did not answer correctly about the original anchor with medium or

high confidence (in which case the program had to explain the answer to the original anchor). These facts indicate that the anchor used was an effective starting place for remediating the primary misconception.

- On the average, 7.5 example situations were needed to bring the subject to a successful conclusion of the session. This supports the need for many nodes in the networks, and indicates that the original anchor and problem were indeed "distant" analogies.

- Most of the instances of giving explanations or strong hints occurred in the area of the molecular model example. This and other observations from the experiment will provide useful and detailed information concerning where students have the least and most difficulty, and where in the example network more bridges need to be added.

- Only one of the 15 subjects spent any time in the area of the network between the book on the spring and the book on the hand, suggesting that the book on the spring would serve equally well as the original anchor. (This finding is in agreement with the results of previous experiments by Clement and Brown).

- Three of the sessions involved pairs of subjects using the tutor. They were asked to discuss their beliefs about the questions and come to an agreement about an answer. These sessions were quite animated. There was a high degree of motivation, verbalization, and serious re-

flection about the problems. Only rarely did the interviewer need to intervene and ask for verbalization of thoughts or reasons for answers. These interactions are encouraging, and we plan to do further studies using pairs.

- Students' comments about the program, elicited at the conclusion of the session, were in general favorable regarding the effectiveness of the learning experience and their enjoyment in using the tutor.

Defining instructional contexts

We are starting to ascertain the instructional contexts in which the strategy is and is not effective. We have noticed four categories that student behavior patterns seem to fall into. The first, for which the strategy is most effective, includes students for whom introducing bridges and encouraging analogical comparisons lead to changing belief to a confident correct understanding of a misconceived situation. They needed few or no hints, and were usually confident in both their correct and incorrect answers. The second group often needed to be given strong hints or explanations of correct answers. This group is perhaps not proficient at analogical thinking, or was not convinced by the particular example situations which they were given. A third group is characterized by not making much progress toward understanding the target problem (the book on the table) until the causal model was introduced, which then led to a complete reversal in their intuitions regarding

the target. It is not clear whether thinking about the previously presented bridging analogies contributed to this change, or whether giving the causal model alone would have been just as effective. The fourth group is characterized by indicating a low confidence for all answers given (both correct and incorrect). Our impression of this group is that they were not mentally engaged in the thought experiments about the physics of the example situations, or had low self esteem.

This grouping is not a partition of the subjects into four distinct groups. Some students exhibited characteristics of more than one group. We have done only preliminary work on how a computer tutor could use the information at its disposal to determine which of these four behavioral contexts a student is in, and what alternate tutoring strategy might be invoked. For example, a student indicating low confidence on most of her answers might occasionally be given positive feedback on correct anchor cases. A student who is only changing her mind as a result of strong hints might be presented with a causal model earlier than she otherwise would. A student for whom the strategy can't find an anchor might be given remedial instruction or a didactic explanation arguing against the misconception. We expect that the context and appropriate strategy will fluctuate within a tutoring session.

In summary, there is evidence that the strategy works effectively for many students. We are encouraged by the levels of cognitive involvement in the thought experiments which the subjects exhibited, especially in con-

trast with the more passive role that students tend to take in traditional explanation oriented methods of teaching, or in computer aided instruction that provides immediate feedback for wrong answers. The formative evaluation provided useful information for the improvement of the specific example network used, for improving the bridging analogies strategy, and for implementing alternative tutoring strategies.

5 Implications for a more intelligent tutoring system

The current implementation of the bridging analogies strategy has served well as a research vehicle. Video taped tests showed that in certain situations the strategy engages students and brings about significant changes in belief. But the tests also showed that the current implementation is not flexible or complete enough to act as a stand-alone tutor (at least in the domain used). The knowledge representation and decision control power of AI technology may provide the leverage needed to design a robust tutor which utilizes bridging analogies as a central strategy.

The tests indicate that alternative strategies, such as presenting causal models and giving leading questions, may be necessary as well. Also, with no model of the student's beliefs, the program had to blindly present example situations until one of the student's "critical points" was reached.[16]

[16] We would like to avoid presenting superfluous examples to the student. But we cannot

Below are three recommendations for more intelligent versions of computer tutors that automate the bridging analogies strategy[17]:

- As a tutoring session progresses, construct a model of the student's beliefs to be used to make tutoring decisions. Data from answers, confidence ratings, reasons for answers, changes of mind, etc. contribute to the student model.

- Select the appropriate bridge between the current problem and anchor (currently done with a lookup table specified by the instructional designer) via an intelligent inference mechanism that takes into account key features of the problem and anchor, and information from the student model.

- Incorporate alternative tutoring strategies, such as using causal models, extreme cases, counter examples, and leading questions, to be used

advocate trying to move directly to critical points (assuming for the moment that this would be possible) because the critical point may represent a cumulative effect from several examples.

[17]Two of us (Murray and Schultz) are in the process of re-designing the tutor for an artificial intelligence workstation using AI programming technologies. The research described in this paper was conducted under the auspices of the Cognitive Processes Research Group at the University of Massachusetts. The re-designed tutor will be written under the auspices of the Intelligent Tutoring Systems Group, in the UMass Department of Computer and Information Science.

in situations where bridging analogies is inappropriate. Invoke these alternate strategies according to production rules which test state variables related to the student model and discourse context.

6 Distinguishing features of this research

As mentioned in the Introduction, our focus on modeling the process of tutoring (specifically, using Socratic type dialogue and carefully chosen example situations) contrasts with many ITS research efforts which focus on domain task analysis and modeling domain expertise. It also contrasts with those projects which focus on designing instructionally effective learning environments or simulations.

Another comparison involves the fact that many existing ITS systems teach skills. Examples are: Shute & Bonar (86) and Streibel et al. (86) (scientific inquiry skills); Brown et. al.'s SOPHIE (82) and Clancey's GUIDON (82) (diagnostic skills); and Anderson et. al.'s LISP tutor (85) (programming skills). Our research is aimed at remediating conceptual difficulties, as opposed to remediating bugs in factual or procedural knowledge. Other ITS science efforts are aimed at teaching conceptual knowledge, but they do so by leveraging the computer's unique ability to simulate physical systems. As mentioned above, we use a discourse strategy to appeal to the logical and/or intuitive sensibilities of the learner. Incorporating semantically rich learning environments in our system at this point would cloud the research

questions, but maximum learning leverage would of course be obtained in a system using both rich environments and intelligent discourse. Also, we foresee other "strong method" tutoring strategies being developed to teach science. Strategies are needed to teach factual knowledge, problem solving skills, and scientific inquiry skills, and to guide students' exploration of learning environments that provide new types of experiences.

Below we summarize what we believe are the distinctive (though not unique) features of this research:

1. Relying on previous cognitive studies of a tutoring strategy.

2. Testing the simulated strategy early on, before large effort is expended in coding an elaborate tutor.

3. Incorporating multiple evaluation-redesign stages using interviewed subjects before a final robust system is built.

4. Being *explicit* about the tutoring strategy used.

5. Researching the strengths and limitations of a highly context sensitive ("strong method") strategy in isolation.

6. Focusing on a single strategy which is ultimately intended to be used in concert with other strategies in an intelligent tutoring system.

7. Focusing on changing students' intuitive beliefs and teaching qualitative conceptual knowledge, as opposed to teaching or debugging skill

knowledge.

8. Using of analogies to extend the locus of applicability of existing intuitions.

9. Our main instructional leverage comes from computer generated Socratic type dialogue, as compared with computer simulation, as is common in other science ITS's.

10. Focusing research on the process of tutorial interaction, rather than on task analysis.

7 Conclusions

The discouraging results of recent broad-based education studies (The National Commission on Excellence in Education, 1983, and The National Science Board Commission on Pre-college Education in Math, Science, and Technology, 1983) have spurred concerns about the quality of science education. Educational researchers are actively studying the process of learning science and are searching for innovative instructional approaches. Some of this research deals with how students' preconceptions influence their learning. Research has shown that misconceptions in the sciences are widespread, debilitating, and resistant to change. Cognitive studies have categorized many of these misconceptions and are experimenting with techniques for dealing with them.

A tutoring strategy found to be useful in remediating physics misconceptions in written explanation and one-on-one tutoring studies has been simulated in a computer tutor, which we call the bridging analogies tutor. Preliminary results from a formative evaluation involving videotaped sessions with the tutor in the domain of Newton's third law are encouraging. The program engaged subjects in thought experiments at an intuitive level, and resulted in changes in belief about misconceived problem situations. The evaluation resulted in many specific implications for computer assisted instruction in the domain used, and for using the bridging analogies strategy in general.

Though conclusions about the general effectiveness and applicability of the bridging analogies strategy must be considered cautiously, we believe that the strategy shows promise for some instructional contexts, and we plan to continue the research. Taking into consideration the results of the evaluation, we have given suggestions for a re-design of the tutor, using artificial intelligence programming technology, to incorporate student modeling, intelligent example selection, and alternate tutoring strategies.

Further work is needed to codify human tutoring strategies, test their effectiveness in computer learning environments, and determine how to orchestrate the different strategies in computer tutoring sessions. We have relied on previous cognitive studies of misconceptions and tutoring strategies in to formulate our ideas. This research exemplifies the importance we

place on such studies in the early stages of designing intelligent computer tutors.

References

[1] Anderson, J. & Reiser, B. (1985). "The Lisp Tutor," BYTE, April 1985.

[2] Anderson, J., Boyle, F., Farrell, R. & Reiser, B. (1984). "Cognitive Principles in the Design of Computer Tutors," Cog. Sci. Soc. proceedings.

[3] Brown, D., & Clement, J. (1987). "Overcoming Misconceptions in Mechanics: A Comparison of Two Example-Based Teaching Strategies," a paper presented at the American Educational Research Association conference, Washington, DC.

[4] Brown, J. S., Burton, R. R., & deKleer, J. (1982). "Pedagogical, Natural Language, and Knowledge Engineering techniques in SOPHIE I, II, and III," in *Intelligent Tutoring Systems*, Sleeman & Brown, Academic Press, New York, NY.

[5] Burton, R. & Brown, J. S. (1982). "An Investigation of Computer Coaching for Informal Learning Activities," *Intelligent Tutoring Systems*, Sleeman & Brown (Eds.), Academic Press, New York, NY.

[6] Camp, C., Clement, J., & Schultz, K. (1987). "Eight Lessons Designed to Overcome Misconceptions in Physics: Relative Motions, Frictional Forces, and Newton's Third Law." Cognitive Processes Research Group technical report, University of Massachusetts.

[7] Clancey, W. J. (1982). "Tutoring Rules for Guiding a Case Method Dialogue," *Intelligent Tutoring Systems*, Sleeman & Brown (Eds.), Academic Press, New York, NY.

[8] Claxton, G. (1985). "Teaching and Acquiring Scientific Knowledge," in *Kelly in the Classroom: Educational Applications of Personal Construct Psychology*, T. Keen & M. Pope (eds.), Cybersystems Inc., Montreal.

[9] Clement, J. (1982). "Students' Preconceptions in Introductory Mechanics," American Journal of Physics, Vol. 50 No. 1.

[10] Clement, J. (1987a). "The Use of Analogies and Anchoring Intuitions to Remediate Misconceptions in Mechanics," a paper presented to the American Educational Research Associntion, April, 1987.

[11] Clement, J. (1897b). "Overcoming Students' Misconceptions in Physics: The Role of Anchoring Intuitions and Analogical Validity," *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*. Ithaca, NY.

[12] Clement, J., & Brown, D. (1984). "Using Analogical Reasoning to Deal With 'Deep' Misconceptions in Physics," appendix to NSF proposal: An investigation of methods for overcoming conceptual difficulties in science, University of Massachusetts, Amherst.

[13] diSessa, A. (1985). "Knowledge in Pieces," Address to the Fifteenth Symposium of the Jean Piaget Society.

[14] Ginsburg, H. & Opper, S. (1979). *Piaget's Theory of Intellectual Development, An Introduction*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

[15] Halloun, I.A., & Hestenes, D. (1985). "The Initial Knowledge State of College Physics Students," *American Journal of Physics* Vol. 53, No. 11.

[16] Lawson, A. & Wollman, W. (1975). "Physics Problems and the Process of Self- Regulation," *The Physics Teacher*, Nov. 1975.

[17] Littman, D, Pinto, J., & Soloway, E. (1886). "An Analysis of Tutorial Reasoning About Programming Bugs," AAAI 86 Proceedings.

[18] McCloskey, M. (1983). "Intuitive Physics," Scientific American, Vol. V48.

[19] McDermott, L. (1984). "Research on Conceptual Understanding in Mechanics," *Physics Today*, July 1984.

[20] Murray, T. & Woolf, B. (1986). "Knowledge Representation in a Physics Tutor," UMass Dept. of Computer and Information Science Technical Report 86-37.

[21] Murray, T., Woolf, B., Brown, D., Clement, J., & Schultz, K. (1986). "Tutoring Strategies for Computer Tutors which Deal with Misconceptions in Physics," working paper.

[22] Murray, T., Clement, J., & Brown, D. (1985). "A Computer Program which Utilizes a 'Bridging Analogies' Tutoring Strategy for Dealing with Misconceptions in Mechanics," Cognitive Processes Research Group Tech. Report.

[23] The National Comission on Excellence in Education (1983). "A Nation at Risk: The Imperative for Educational Reform."

[24] The National Science Board Commision on Pre-college Education in Math, Science, and Technology (1983). "Educating Americans for the 21st Century."

[25] Schultz, K., Murray, T., Clement, J., & Brown, D. (1987). "Overcoming Misconceptions with a Computer-Based Tutor," *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*. Ithaca, NY.

[26] Shute, V. & Bonar, J. (1986). "Intelligent Tutoring Systems for Scientific Inquiry Skills," Proceedings of the Cognitive Science Society.

[27] Stevens, A. & Collins, A. (1977). "The Coal Structure of a Socratic Tutor," BBN Report No. 3518.

[28] Streibel, M., Stewart, J., Koedinger, K., Collins, A., Jungck, J. (1987). "MENDEL: An Intelligent Computer Tutoring System for Genetics Problem Solving, Conjecturing, and Understanding," *Machine-Mediated Learning*, Vol. 2. No 1 & 2, Friedman & Howard (Eds.).

[29] Tennyson, R. & Park, O. (1980). "The Teaching of Concepts: A Review of Instructional Design Research Literature," *Review of Educational Research*, Vol. 50 No. 1.

[30] Tversky, A. & Kahneman, D. (1974). "Judgement Under Uncertainty: Heuristics and Biases," *Science*, Vol. 185.

[31] Van Lehn, K. (1983). "Felicity Conditions for Human Skill Acquisition: Validating an AI-Based Theory," Xerox Palo Alto Research Center Technical Report CIS-21.

[32] White, B. & Frederikesen, J. (1986). "Intelligent Tutoring Systems Based Upon Qualitative Model Evolution," AAAI Conference Proceedings.

[33] Woolf, B. (1984). "Context Dependent Planning in a Machine Tutor," UMass Dept. of Computer and Information Science Technical Report 84-21.

Figure 4: An Algorithm for the Bridging Analogies Strategy

APPENDICES

A Program data structures and algorithm

Domain information is stored in text files, making the program which drives the tutor independent of domain information. There are three domain files. The Topic File determines the sequence of topics to be presented (this paper discusses only one of the topics), and specifies which of the examples in the topic's network will serve as the original anchor and original problem situations. The Network File contains triplets of numbers, each number representing an example situation node. Each triplet defines a unique bridge between two other nodes, and defines the network as in Figure 1. The table thus defined is searched when the program bridges between a problem and an anchor. "Wild cards" are incorporated, allowing the designer to specify the use of a certain bridge whenever a specified node is the problem (or anchor), regardless of the anchor (or problem).

The Situation File contains text for the example descriptions, detailed example descriptions, questions, multiple choice answers, correct answer, and a hint or explanation for the correct answer. This information is organized using keywords, and is parsed and loaded into PASCAL record structures when the program is executed. These record structures also keep track of the student answers and confidence ratings for each node.

The bridging analogies strategy is a doubly recursive algorithm for traversing the network (the network is topologically equivalent to a binary tree). The algorithm has been described operationally in Section 3. It appears in pseudo- code in Figure 4. The code shows what parameters are used to make decisions in the procedure and where in the procedure the recursion occurs.

The procedure Teach-Topic is called iteratively for each topic in the Topic File. The Bridge procedure contains the algorithm for traversing the network. The program is capable of "finding" the original anchor, but this capability was switched off (and the original anchor is given in the Topic file, as described above) in the test runs we report in this paper. It finds an original anchor by presenting the student with all the situations linked to the original anchor in order of increasing simplicity, until the student answers one correctly with high confidence. For large networks this may

be more efficient than assuming that the best original anchor is the very simplest example.

The algorithm is relatively simple, yet due to its recursive nature, and the fact that the confidence level is used differently in sections 2 and 3 of the Bridge procedure, it is quite difficult for a student using it to distinguish a predictable pattern. The presentation of examples, hints, and comparisons seems random to one not familiar with the algorithm. Recall that we do not explicitly tell students whether they are correct on any questions, so it is to great advantage that the student cannot (or rarely has in our tests thus far) determine the correctness of an answer merely from the type of action taken by the tutor.

Figure 5: Sample section of a Trace file

Figure 6: Diagramatic Representation of the Program Trace

B Data Analysis

There were three sources of data: notes taken by the interviewer during the sessions, a computer trace of the tutor's actions and user responses, and videotaped recordings of the sessions. Analysis of the videotapes is described in Schultz et. al. (87). Here we will outline our method for analyzing the computer traces. The trace records all important events in a readable form, as in Figure 5. This information was transferred (by hand) to a diagramatic representation, as in Figure 6.[18] Circled nodes are those visited. Next to each circled node the quality of the answer (ok or wrong) and the confidence rating (from 1-blind guess to 5–I'm sure) are recorded. Single line arrows go from problems or anchors to the next bridge. Double lined arrows go from bridges to problems, and show where the subjects changed their mind from wrong (or right with low confidence) to correct (or medium to high confidence) for the problem pointed to. Circular lined arrows which connect a node with itself represent changing their mind on the question most recently asked, or, if an "H" is shown, they were given a hint and asked the question again (as would happen if no bridge were available between two nodes). Not represented in the diagram, but easily inferred from the information in it, are the places where the tutor asks for a comparison between the most recently given situation and a previous one. Recall that after each such comparison the student is asked if she wants to change her mind on either or both situations.[19]

Data analysis is greatly enhanced by the diagramatic representation. The subjects' path through the example network, their changes of mind, and where hints were needed, are highly visible. We can easily see regions of the net where the subjects needed the most bridges. From this we can infer which attributes of the original problem contributed to their misconception.

For each subject's network, the following data were determined: the

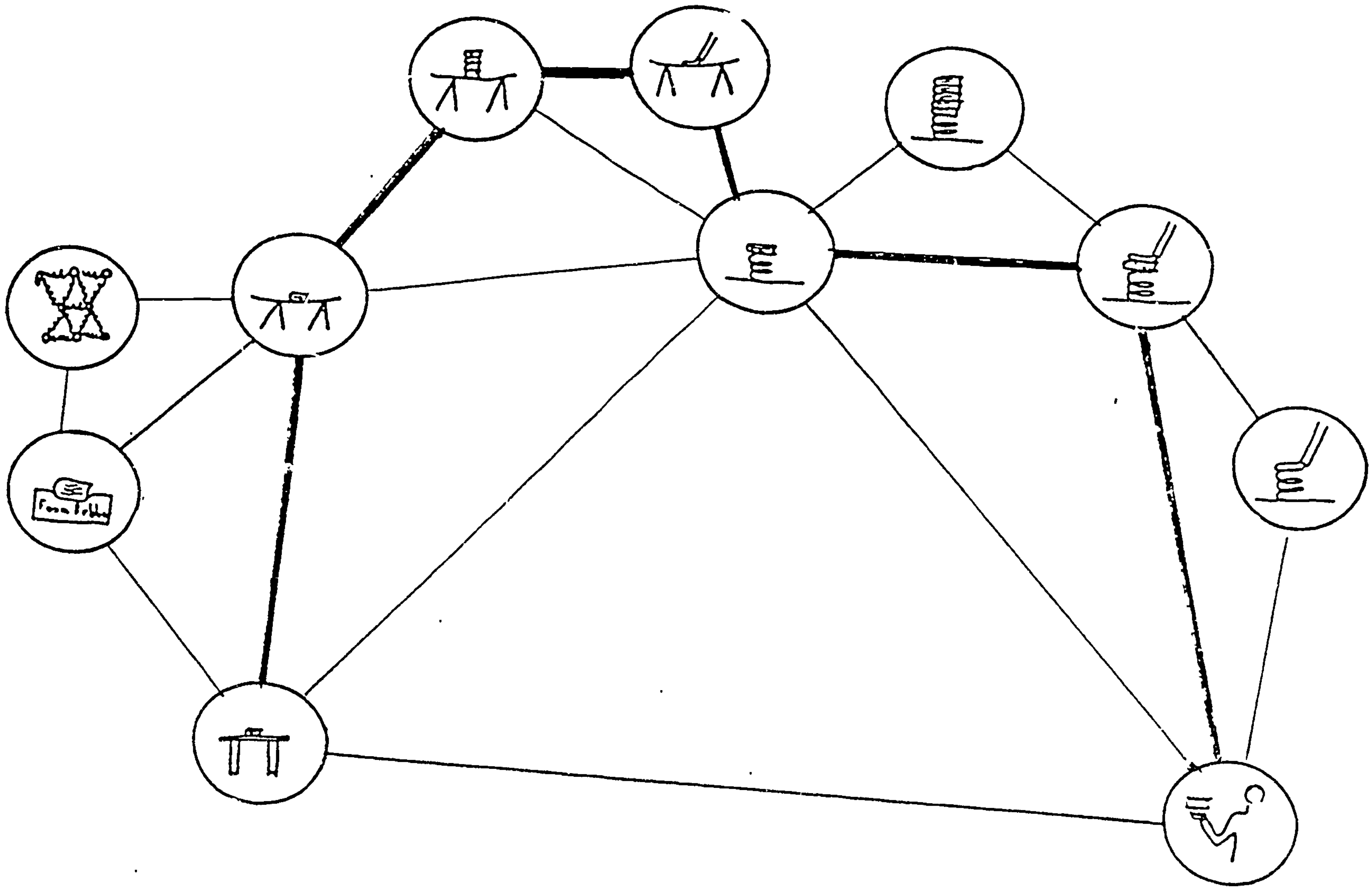
[18] This figure shows the entire network used. The portion to the left, between the book on the table and a fly on the road, is ignored in other parts of the paper for simplicity.

[19] With some programming effort (most of it for graphics) the conversion from trace to diagram could be automated, but we chose not to do so.

number of nodes visited, the number of hints given, the number of changes in mind, the total change in the original problem from the beginning to the end of the session, the "critical" changes in mind (defined below), and the "focal regions" (defined below). A summary of this information is given in the Results section of this paper.

The network was divided into four "focal regions," each addressing a different key attribute of the target problem. The "focal regions" data specifies in which regions the subject visited two or more nodes.

"Critical" changes give a more reasonable estimate of the student's thinking than the total changes of mind, and are defined as follows. Often, one realization on the part of the subject leads to changes of mind for several situations. For example, a student who is being shown a hand on a table situation (node 10) is four recursive levels deep in the algorithm, having answered incorrectly questions about a book on a table, a book on a flexible board, and many books on a flexible board. Considering the hand on the board may be a critical analogy for this student. After answering it correctly he may realize that several of his previous ansers should be changed. As the recursive algorithm unwinds (see Figure 4), he would be given three comparison questions in a row, and opt to change his previous answer for each of these to correct with high confidence. When the subject changes his mind in a "chain" like this, all changes can be the result of a single change in belief. "Critical changes" are the first changes in any chains of two of more changes, plus all the changes that are not in a chain.



Network of example situations for existence of normal forces

Figure 1

Imagine a medium size textbook resting on a dining room table.

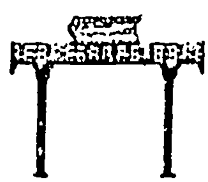
While the book is resting there the table:
A. IS exerting a force up on the book
B IS NOT exerting a force up on the book

Please rate your confidence in this answer:
'blind guess' 'not very conf' 'somewhat conf' 'fairly conf' 'I'm sure'

Imagine that you are holding a textbook in your hand.

While the book is resting there your hand:
A. IS exerting a force up on the book
B. IS NOT exerting a force up on the book

(*** again give an answer and confidence ***)



For "a book in your hand," you said:
your hand IS exerting a force up on the book
(with high confidence).

But for "the book on the table," you said.
the table IS NOT exerting a force up on the book
(with fair confidence).

Explain why these answers are different.

*** Do you want to change your mind on either or both of them? ***

Figure 2: A sample dialogue showing the original problem and anchor

The Bridging Analogies Strategy

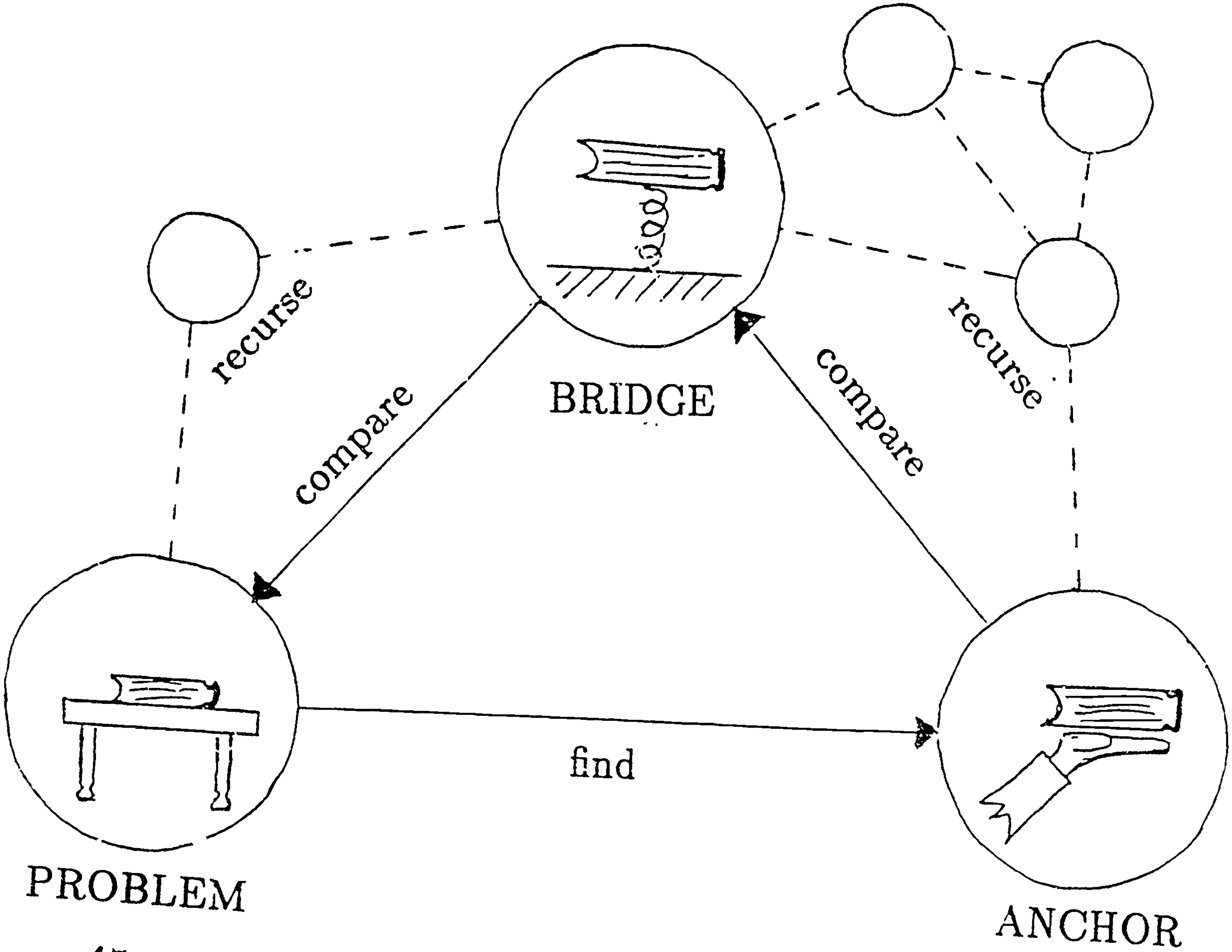


FIGURE 3

Figure 4

```
Procedure Teach-Topic (original-problem, original-anchor)

   Describe-and-question(original-problem)

   Describe-and-question(original-anchor)

   Compare-and-allow-change(original-problem. original-anchor)

   IF anchor wrong or confidence<3
      THEN Hint(original-anchor)

   IF problem is right with confidence>2
      THEN RETURN {no misconception}
      ELSE Bridge(original-problem, original-anchor)


Procedure Bridge (problem, anchor)

   {The anchor was previously answered correctly with
    confidence >2 and the problem was previously answered
    incorrectly or with confidence <3 }

   {1. find the bridge}

        Find-a-bridge-between(problem, anchor)
           IF no bridge exists
              THEN Hint(problem) AND RETURN
           Describe-and-Question(bridge)

   {2. establish bridge-to-anchor analogy:}

           IF bridge wrong or confidence<4
              THEN Compare-and-allow-change(bridge, anchor)

           IF anchor changed to wrong or confidence<3
              THEN Hint(anchor)

           IF bridge wrong or confidence<3
              THEN bridge(bridge, anchor)    {first recursive call}

   {3. establish problem-to-bridge analogy:}
```

(cont.)

Figure 4 (cont)

```
Compare-and-allow-change(problem, bridge)

IF problem is wrong or confidence<3
   THEN bridge(problem, bridge) {second recursive call}

{4. At this point the problem should have been answered correctly}

RETURN
```

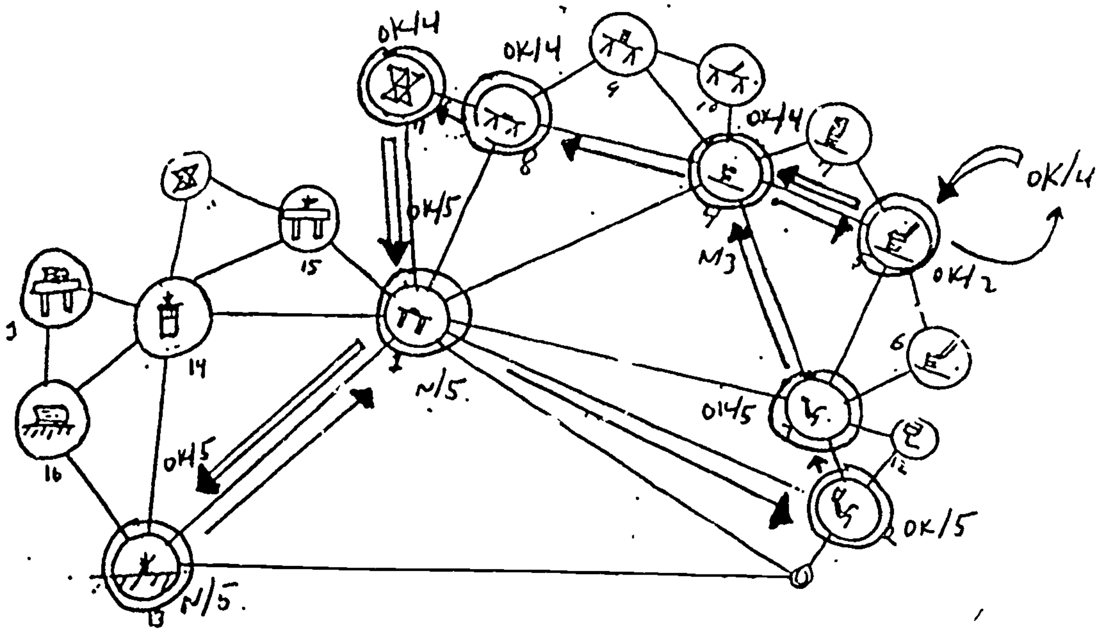
```
Procedure Hint (an-example)

   {iteratively gives the hint and re-asks the example's
   question until the student answers correctly, or until
   given three hints in a row, in which case the student
   is told exactly what to enter as the answer(!)}
```

(Note: Confidence ratings from 1 to 5 are for 'blind guess' to 'I'm sure' as in Figure 2.)

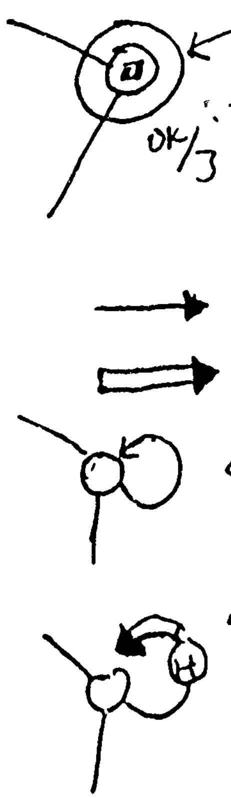
Figure 5: Sample section of a trace file

```
START SESSION:
program version: 11/18/86.1
Name:  Willard O. Oz
Session ID:  292

Topic: BOOKTABLE
O-prob, O-anch, O-sub:  13  0  8

....
....

ASK  1  book on the table
Response: B (not OK) Confidence: 5

ASK 2 many books in hand
Response: A (OK) Confidence: 5

*BRIDGE: P 1 A 2 B 3   11:50:35
ASK 3  one book in hand
Response: A (OK) Confidence: 5

COMPARISON: 2 and 1; no change

*BRIDGE: P A B  11:50:47
ASK 4 book on spring
Response: b (Not OK) Confidence: 3

Comparison: 4 and 3; no change

*BRIDGE: P 4 A 3 B 5 11:51:17
ASK 5 hand on  book on  spring
Response: A (OK) Confidence: 2

Comparison: 5 and 3; change 5

ASK 5 hand on book on spring
Response: A (OK) Confidence: 4

Comparison: 5 and 4; change 4

ASK 4 book on spring
Response: A (OK) Confidence: 4

Comparison: 4 and 1; no change

*BRIDGE: P 1 A 4 B 8  11:53:03
ASK 8 book on board
Response: A (OK) Confidence: 4

....
....
```

Figure 6 Diagramatic representation of the program trace



KEY



visited

answer quality

confidence rating (1 to 5)

to next bridge

changed mind from wrong to right

changed mind on questions just asked

hint given