Varieties of Learning from Problem Solving Experience*

Juliana Lancaster
BDM Corporation
1300 N. 17th St., Suite 950
Arlington, VA 22209

Janet L. Kolodner
School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332

March 30, 1988

1 Introduction

Problem solving ability appears to be acquired in two ways. Some of the knowledge required for problem solving can be obtained through direct instruction or "book learning", but major portions of what experts know is derived directly from their experiences with problems in the domain of interest. As work in artificial intelligence has begun to focus on highly complex domains, understanding the contribution of experience to knowledge and behavior has become increasingly important.

As a result, researchers have begun to consider how experience leads to new knowledge and, specifically, the role of single cases in learning. In earlier work on psychiatric diagnosis, Kolodner (1982) developed some proposals about how an individual problem solving experience can lead to changes in the knowledge base of the problem solver. That research focussed on the changes in knowledge that can result from a diagnosis that was initially incorrect and then corrected. We argued that an analysis of the knowledge used in making a new, correct diagnosis would lead to the individual becoming aware of the errors or missing knowledge in his knowledge base that led to the original incorrect diagnosis. At that point, the problem solver would modify his knowledge appropriately to ensure that the same error would not recur. The modification process had two parts to it: modification of the problem solver's general knowledge and problem solving procedures and indexing of the experience in memory so that it could be remembered during later problem solving.

Our goal in the study reported here was to look at less experienced problem solvers to see what processes were involved in their learning. The processes involved in modifying general knowledge and problem solving procedures requires a fair amount of expertise on the part of the problem solver, both about the domain and about what could easily be done wrong in problem solving. Novices know very little about either of these things to do a full analysis themselves of their actions. We wanted to find out what kinds of learning procedures are available to them, what they can learn through problem solving experience, and what kinds of supervision are necessary for learning to happen. We also looked at which of their experiences they remembered to help them during later problem solving.

The domain we have been exploring is automotive diagnosis. This domain is of interest for several reasons. The automobile engine is a highly complex entity, consisting of several interacting subsystems. Failures in any component or subsystem produce symptomatic behaviors in engine performance, but the symptoms are not necessarily unique to a given failure. In other words, several failures can produce the same symptom(s). It is the mechanic's task to identify the correct failure.

*This research was supported in part by the Army Research Institute for the Behavioral Sciences under Contract No. MDA-903-86-C-173. Thanks to Phyllis Koton, Joel Martin, and Michael Redmond for commenting on an earlier draft of the paper.

Furthermore, the domain is knowledge-rich and both depth of knowledge and the ability to use it are important in reaching a good diagnosis. A student can be taught about cars and troubleshooting in general, but since cars vary extensively, and models may change significantly from year to year, no one can learn everything within an instructional setting. Clearly experience with different types and ages of cars is necessary for an individual to develop adequate expertise. An expert mechanic must have sufficient experience to allow him to draw his own generalizations about cars and to organize and access his knowledge in a manner appropriate to the specific car and problem he is considering.

In the work reported here, we observed the diagnostic process in several individuals and compared the knowledge used to solve successive similar cases. We also observed the interactions between problem solvers (students) and outside sources (books and instructors). The analysis presented here uses data presented in Lancaster & Kolodner (1987). That work identified the types of knowledge and diagnostic processes used by novice car mechanics.

Four automotive mechanics students at a local two year technical school were observed while diagnosing car failures. Six problems were presented at weekly intervals and think-aloud protocols were collected while the students worked. Each week, after all students had attempted to solve the problem, the instructor demonstrated the correct diagnostic procedures and his method of identifying the failure. Thus each student had a opportunity for feedback about his performance regardless of whether or not he diagnosed the problem correctly. Each failure was introduced into a car deliberately and each failure was caused by only one failed part. Four of the six problems presented the same symptoms initially, and a fifth problem presented a related symptom.

After all protocols were coded, two sequential protocols for each student were selected. The two protocols were selected to have the same initial symptom, but different failures. For each of these protocols, a description of the knowledge used and sought by the student during diagnosis was compiled. The descriptions of the two protocols were then compared and differences were noted. Of particular interest were occasions in which the student appeared to use knowledge in solving the second problem that he did not have in solving the first problem. Changes in the knowledge used showed us what students learned between solving the two cases. We then examined the protocols and the instructor's explanation of how to solve the first problem to see if we could identify how these items were learned. Sometimes this was obvious – a student had asked a question previously or the instructor had presented the relevant material in a previous explanation. Sometimes it was not obvious – what was learned might have been presented in class between the two exercises. We focussed on those learned items where we could identify the situation in which the new concepts were learned.

2 What was learned

As reported in our earlier paper (Lancaster & Kolodner, 1987), all of the students seemed to possess the same types of knowledge structures: a causal model of the engine, symptom-fault sets linking particular symptoms to the failures they might indicate, and a collection of troubleshooting guidelines and procedures. However, the students at various levels of training differed in both the amount of knowledge they had and the organization of that knowledge. The novice's knowledge was unorganized, sparse, and sometimes wrong. The intermediate student had a better organization on his knowledge (e.g., he grouped parts in the same system together) and some misconceptions had been corrected, but it was still incomplete. The advanced student had a much more complete knowledge base, and many of the misconceptions of the novice and intermediate students were corrected. In particular, the advanced student knew more about what failures normally look like than did the intermediate student, who knew more about how the car was supposed to work than about its malfunctions. The novice knew almost nothing about malfunctions.

The novice student, having learned the rudiments of how a car works in classroom instruction, was primarily focussed on acquiring the necessary organization of the causal model to allow systematic tracking through the engine in search of a failure. For example, diagnosing a fuel system restriction in the earlier problem, the novice moved directly from the symptom (car cranks but will not start) to the hypothesis that it was not getting enough fuel and concluded that the fuel line was restricted. In a later case with the same symptom, the novice student instead showed that he had subsumed the fuel restriction hypothesis under a more global hypothesis of fuel system failure. He first checked the end point of the fuel system for evidence of failure and then began checking components within the system. Thus he had modified his knowledge base in an important way between cases. He had collected the fuel system components into a system, allowing him to consider the system as a whole as his first hypothesised failure source. As a consequence, his trouble- shooting behavior became more systematic as he only considered specific components of a system if the system as a whole was shown to be the source of the failure. This type of learning can be termed *knowledge reorganization.*

The novice also improved his problem solving skills. In the first case, he did not take action to confirm his diagnosis. Rather he was content to accept the hypothesis on the basis of its being a possible failure matched to the symptom. After attempting to solve the problem himself, he listened to the instructor's explanation of how to diagnose the a car's problem. The instructor, as part of his explanation, showed that it is necessary to confirm any hypotheses that are made. In the second case he solved, the novice had learned that confirming hypotheses is a necessary component of problem solving. He was thus also refining his problem solving skills.

In contrast, the intermediate student, having established the basic hierarchical organization of his causal model and symptom-fault sets, was primarily engaged in adding new information to his knowledge bases. Specifically, he showed evidence of adding to the faults associated with a given symptom in his symptom fault sets, learning new procedures for diagnosis, and identifying the locations and functions of new components. For example, on the earlier problem, the intermediate student made a protracted search for the fuel pump relay. On the later case, he was able to locate and test the relay at the appropriate time with ease. Thus, we see that he had learned not only its location but also how it was connected within the fuel system. Similarly, he employed techniques on the later problem that he had apparently not known when solving the earlier problem, leading to more efficient testing of hypotheses and an increased likelihood of getting correct, useful information from his testing.

The advanced student engaged in yet another type of learning that we have termed *refinement.* He already knew most of the information needed to successfully solve every problem presented. The changes in his knowlege between cases reflected primarily a reordering of faults in the symptom-faults sets to reflect new probablilties of occurrance. In other words, after solving a case in which the correct hypothesis was one he considered late in his troubleshooting procedures, the advanced student returned to that hypothesis earlier in his sequence on the later problem. He showed no notable changes in the overall organization or content of his knowledge between problems.

In addition, the advanced student seemed to have an additional model available to him that the intermediate and novice students were missing: a model of a *malfunctioning* car. Both the novice and intermediate student seemed to diagnose faults by first zeroing in on a system that might be faulty and then checking the behavior of the car against the expected behavior of a *working* car. What differentiated the novice's behavior from that of the intermediate student was that the novice might focus on a particular part, while the intermediate student focused on a system. The more advanced student (and the instructor), however, seemed to know what to expect in a *malfunctioning* car. In other words, he not only had a *working model* of the car, but he also had a *malfunction model* of the car. We can see this in his behavior when the testing equipment was faulty. The advanced student was the only one who could recognize that the readings he was getting from the test equipment were faulty. The intermediate student was unable to differentiate a bad reading on test equipment from a faulty car.

3 Learning Processes

Several learning processes were identified by examining the protocols:

Learning by understanding explanations: After the students had diagnosed each problem, the instructor demonstrated the correct, or optimal, diagnostic path for the problem. In doing so, he enumerated both the reasons for considering each hypothesis and the diagnostic and test procedures he was using. The students attempted to modify their knowledge to match the instructor's wherever his procedures, hypotheses, or explanations differed from theirs. We have termed this process *learning by understanding explanations* (Redmond & Martin, 1988, Martin & Redmond, 1988). It is a combination of learning by observing an expert (Mitchell, et al., 1985) and learning by being told. In learning by understanding, the learner has an opportunity to observe another person (in this study, the instructor) solving the same case while providing an ongoing explanation of the knowledge and processes used. The learner notices those points at which the instructor's knowledge differs from his own and modifies his knowledge to bring it into agreement with the instructor's. There are a variety of things that can be learned by this method: new organizational structures, as when the novice collected the fuel system components into a system; new diagnostic strategies, as when the novice learned to test system endpoints or when the intermediate student learned how to use trouble trees; new causal connections in the causal model; new test procedures; and refinements of any of these things.

A more complete explanation of this process can be found in (Redmond & Martin, 1988) or (Martin & Redmond, 1988). In short, a student learning by this method attempts to explain each of the things the teacher is doing, and then uses the teacher's additional comments and explanations to bridge the gaps in his understanding. Thus, if the teacher proposes a particular hypothesis given a particular set of symptoms, the student (internally) attempts to explain why that is a good hypothesis. The student may not know why a particular hypothesis is appropriate, but the teacher's explanation allows him to fill in the gaps in his attempted explanation. Without the teacher's explanations, the best the student can do is to draw associations.[1]

Active gap filling: While engaged in diagnosis, a student would sometimes realise that he did not know a specific piece of knowledge about the behavior of a component, or the connections between a component and the rest of the engine that was needed to solve the current problem. For example, his diagnostic procedures might have told him that he needed to check the input to some component, but he might not have known what the input source was. To find the missing knowledge, he would consult an outside reference source, such as a book or a more experienced individual. The information thus obtained was then incorporated into his knowledge base and was available during later diagnostic sessions. This is the process by which the intermediate student learned the location of and connections to the fuel pump relay. While he knew that he needed to find the fuel pump relay, he was unfamiliar with the particular car he was diagnosing. He knew that manuals provided this kind of information and went to the manual to find out. He went directly to the appropriate schematic in the manual and then spent considerable time using that schematic as a map of the engine and eventually finding the part he was looking for. In other words, he used an outside source (in this case, the manual) to find out exactly the specific piece of information he was missing.

We have termed this learning process *active gap filling* in recognition of the intentional nature of the learning. To engage in active gap filling, the individual must have enough knowledge about the system, enough problem solving knowledge to know what he ought to be doing, and an adequate organizational structure to allow him to recognize where the gaps in his knowledge might be, or at least to realize when an apparent dead end in his diagnosis might be the result of

[1] This process may remind people of Mitchell's (1985) learning by observing an expert, built into the LEAP system. Learning by understanding explanations assumes that the learner has incomplete knowledge and that the task of the learner is to augment its knowledge. LEAP's method, as well as other methods employing EBL and EBG (DeJong & Mooney, 1986, Mitchell, et al., 1986), assume that the learner has complete knowledge, and allow the learner to better package that knowledge.

a lack of easily obtained information rather than something more extensive. Possibly because of this requirement for a fairly complete and well organized knowledge base, active gap filling was primarily used by the intermediate and advanced students as a learning tool. Our protocols show that knowledge learned through active gap filling is remembered and available for use in later problem solving.

Learning from interpreting feedback: When a diagnostic procedure was not yielding the results the student expected, or when he could not interpret the results from a test, students had to ask for help. In essense, learning from interpreting feedback is a combination of active gap filling and learning by understanding explanations. As in learning by active gap filling, the student is aware of a gap in his knowledge. He may not, however, know what that gap is. And, while active gap filling is usually a process of finding out about some feature of an object, learning by interpreting feedback focuses on procedures: in particular, "what went wrong" with a particular procedure. Learning by interpreting feedback can result in correcting a faulty causal model, but more often involves correcting and refining knowledge about how to do things. This process is similar to learning by understanding explanations, since the instructor may provide a causal explanation of some phenomenon or offer advice about carrying out procedures. It is a more active and goal-directed process, however, initiated by some complication the student is experiencing while solving a problem.

Learning by interpreting feedback is invoked when the student cannot interpret the unexpected results of his problem solving. The process of explaining those results can be done by the student or by asking the instructor. When an intermediate student found that he was getting test results that he could not interpret he went to the instructor for help. On the same problem, the advanced student figured out for himself that the test equipment was faulty, and was able to learn on his own which ranges of test results predicted faulty equipment. The advanced student could do this, while the intermediate one could not, because he had knowledge telling him what things look like when they malfunction. This experience told him what malfunction of a particular instrument looks like. The intermediate student, on the other hand, did not know enough to hypothesize by himself that the test equipment might be faulty. On another occasion, however, the advanced student needed the instructor to provide an explanation to him. He was trying to energize the fuel pump from the battery using a test light to connect them together. He could not get the fuel pump to go on. He, like the intermediate student in the previous example, knew what the results should look like (in this case, he thought the fuel pump should go on) but did not know why the results he was getting were different. He asked the instructor what he was doing wrong. The instructor told him that he needed to use a lead (a wire with no bulb attached) to energize the fuel pump, since with a light attached, the bulb consumes the power from the energy before it gets to the pump. In this case, the advanced student's causal model of electricity was faulty.

Abstraction: Another learning process that was noted was *abstraction*. It seemed that any new information acquired by a student between problems was incorporated into the knowledge base at several different levels. We can see this most clearly in the novice student's behavior. Recall that in one problem, he tracked the failure to the fuel pump but did not examine the system further to distinguish whether the problem was in the fuel pump or in the input to the fuel pump. The instructor followed that session with an explanation of how to diagnose the fuel pump problem. In his explanation, he stated that the endpoints of the fuel system had to be checked for evidence of a problem. The student learned this (by understanding the instructor's explanation) and in the next problem applied it. The student also apparently learned the abstract principle the instructor was illustrating: that system endpoints need to be checked for problems when diagnosing any system. Abstraction is sometimes done spontaneously, as in the example just given, and is sometimes induced by the instructor during explanation. For example, in one instance the instructor walked students through an engine they hadn't seen before, and stated that parts with a particular function had to be found. He then explained what those parts look like in general. This abstract knowledge allows the student to identify the part no matter what type of car he is looking at.

Case-Based Reasoning: A final process that facilitated learned was *case-based reasoning*. In case-based reasoning, a previous case that a problem solver is reminded of provides an answer to a new problem or focuses him on the knowledge needed to solve the problem (Kolodner & Simpson, 1984, Hammond, 1986). While in the previous paragraphs, we have referred to learning processes that allow a problem solver to learn new facts, case-based reasoning is a problem solving method that allows a problem solver to improve its performance without full understanding of the facts. Remembering previous cases allows a problem solver to solve a new problem better than an old one even when the problem solver is missing a causal explanation of why the previous solution did or did not work. Our hypotheses about case-based reasoning in experts predict that those cases that are different than what is expected are the ones that are remembered (Kolodner, 1982, 1983, Kolodner & Simpson, 1984, Schank, 1982). But novices don't always know the norms, so they can't recognize that something is different. The data we collected in this study show three situations in which novices stored cases for later problem solving. First, if a case serves as a strong justification for doing a procedure, then the case was kept and referred to later. The advanced student, in an early problem, spent 45 minutes trying to solve a problem that he could have solved instantly with visual inspection. In the next problem, he did a visual inspection immediately, saying "what we want to do is check the wire...like we did the last time". Second, cases were maintained in memory if they served as an example of a particularly complex or unusual situation. This includes, among other things, cases where a set of symptoms predict a highly unusual fault and cases where some set of symptoms predict a fairly commonplace fault that is hard to diagnose. Both students who solved problem 3 (one intermediate and one advanced) , in which a rocker arm had been removed, for example, were able to do it on the basis of remembering a previous case in which a rocker arm had been broken. A broken rocker arm is a highly unusual problem. In almost all cases, rocker arms outlive the cars they are in. The other students, who had never seen such a problem, were unable to solve the problem (one intermediate and one novice). These two students were able to determine that none of the possible common causes of the symptoms were responsible for the failure of the car, but they were not able to pursue their investigations beyond that point. The two successful students were reminded (based on the particular sound of the car) early in their diagnosis of two cars they had worked with in the previous quarter that had the same problem. This reminding led both of them to try the diagnostic procedures that would lead them to the correct diagnosis and, thus, to the diagnosis. A third instance in which a previous case was kept was when it illustrated something that was not known previous to that case. The case was remembered until what was learned from it was confirmed by a later case. While one of the other procedures might have been used to learn a general concept, the case in which it was learned seemed to remain available to the students until the new item of information was confirmed by another case. We could not determine, however, whether the general knowledge or the case was accessed first in later problem solving.

4 Discussion

While most researchers studying learning have been studying unsupervised learning or learning where the teacher gives an example but no explanations, our observations show that much of the learning that goes on early in problem solving depends on a teacher to give an explanation of the procedures for solving a problem and the knowledge needed to solve it. The causal model describing how a car functions properly, the one describing what malfunctions look like, associations between symptoms and faults, and the problem solving procedures the instructor finds useful motivate his explanations.

Early in learning, students don't know enough to be able to learn exclusively from their own problem solving experiences. For example, the novice who tracks down the error to the fuel pump when it is the fuel pump relay that is faulty (i.e., the fuel pump's imputs are causing it to malfunction, not the pump itself) can learn only if he is given extensive opportunity to attempt to fix the problem or if a teacher intervenes to show him what he was doing wrong. And, when the student is a rank novice, the experience of trying to fix it himself may be so overwhelming that, in the worst case, nothing is learned, while in the best case, learning takes a long time. It might take several attempts, for example, trying to put in a new fuel pump before the student will start thinking that something else is the matter, and at that point, there are so many things that could have gone wrong during his previous attempts (e.g., the new fuel pump may be faulty, he may have connected it up wrong in a whole range of different ways) that he may only be able to track down the problem if he is lucky in his initial guesses.

Nor did our intermediate student seem capable of learning without supervision. While he didn't need help with simple diagnostic procedures, he was unable to differentiate between instrument readings produced by faulty parts and instrument readings produced by faulty instruments. In other words, his missing knowledge about what faults look like required that an instructor intervene to instruct him on that subject.

Not only did students need to know how things work (i.e., have a nearly complete *working model*) to learn in a completely unsupervised setting, but they also need to know how things can go wrong and what things are in the normal realm of possibility. The advanced student, who was by no means an expert, was able to do quite a bit more learning by himself because he knew both of these things. He also, however, needed help from time to time when the knowledge he needed to solve a problem by himself was missing or faulty.

We have attempted to outline the real-world learning procedures students use in learning a complex diagnostic task and the role of the teacher in learning such tasks. While much research has gone into unsupervised learning, little research has focused on processes by which a student learns from a teacher. Among the learning processes we have identified, three require extensive interaction with a teacher: learning by understanding explanations, active gap filling, and learning by interpreting feedback. The particular interactions depend on the knowledge the student already has. More research in this area is surely needed if we want to develop better teaching technologies and practices.

5 References

1. DeJong, G. & Mooney, R. (1986). Explanation-based learning: An alternative view. *Machine Learning*, Vol. 1.

2. Hammond, K. (1986). *Case-Based Reasoning: An integrated theory of planning,learning and memory.* Ph.D. Thesis. Yale University.

3. Kolodner, J. L. (1982). The Role of Experience in Development of Expertise. *Proceedings of AAAI-82.*

4. Kolodner, J. L. (1983). Maintaining Memory Organization in a Dynamic Memory. *Cognitive Science*, Vol. 7.

5. Kolodner, J. L. & Simpson, R. L. (1984). Experience and Problem Solving: A Framework. *Proceedings of the Sixth Annual Conference of the Cognitive Science Society.*

6. Lancaster, J. S. & Kolodner, J. L. (1987). Problem solving in a natural task as a function of experience. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society".*

7. Martin, J. D. & Redmond, M. (1988). The Use of Explanations for Completing and Correcting Causal Models. Submitted to the 1988 Cognitive Science Conference.

8. Mitchell, T. M., Kellar, R. M. & Kedar-Cabelli, S. T. (1986). Explanation based learning: An unifying view. *Machine Learning*, Vol. 1.

9. Mitchell, T. M., Mahadevan, S., & Steinberg, L. I. (1985). LEAP: A Learning Apprentice for VLSI Design. *Proceedings of IJCAI-85.*

10. Redmond, M. & Martin, J. (1988). Learning by Understanding Explanations. *Proceedings of the 26th Annual Conference of the Southeast Region ACM*, Mobile, Alabama.