(2)

AD-A222 428

RULE ACQUISITION EVENTS IN THE DISCOVERY OF PROBLEM SOLVING STRATEGIES

Technical Report PCG - 17

Kurt VanLehn
Departments of Psychology & Computer Science

DEPARTMENT
of
PSYCHOLOGY

Carnegie Mellon University

90 06 05 057

RULE ACQUISITION EVENTS IN THE
DISCOVERY OF PROBLEM SOLVING STRATEGIES

Technical Report PCG - 17

Kurt VanLehn
Departments of Psychology & Computer Science

Carnegie Mellon University
Pittsburgh, PA 15213 U.S.A.

3 March 1989

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT
	Approved for public release;
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
PCG - 17	Same as Performing Organization

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Carnegie Mellon University		Personnel and Training Research Office of Naval Research (Code 1142PT)

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Department of Psychology Pittsburgh, Pennsylvania 15213	800 N. Quincy Street Arlington, VA 22217-5000

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
ame as Monitoring Organization		N00014-88-K-0086

8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO
	N/A	N/A	N/A	N/A

11. TITLE (Include Security Classification)

Rule acquisition events in the discovery of problem solving strategies

12. PERSONAL AUTHOR(S) VanLehn, Kurt Alan

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM 88Jan01 TO 90DEC31	1989 July	57

16. SUPPLEMENTARY NOTATION To appear in Cognitive Science

17	COSATI CODES		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	learning, skill acquisition, problem solving.

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

Although there are many machine learning programs that can acquire new problem solving strategies, we do not know exactly how their processes will manifest themselves in human behavior, if at all. In order to find out, a line-by-line protocol analysis was conducted of a subject discovering problem solving strategies. A model was developed that could explain 96% of the lines in the protocol. On this analysis, the subject's learning was confined to 11 rule acquisition events, wherein she temporarily abandoned her normal problem solving and focused on improving her strategic knowledge. Further analysis showed that: (1) Not all rule acquisition events are triggered by impasses. (2) Rules are acquired gradually, both because of competition between new and old rules, and because of the subject's apparently deliberate policy of gradual generalization. (3) This subject took a scientific approach to strategy discovery, even planning and conducting small experiments.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Susan Chipman	(202) 696-4322	1142 PT

DD FORM 1473, 84 MAR   83 APR edition may be used until exhausted.   SECURITY CLASSIFICATION OF THIS PAGE
All other editions are obsolete.

Table of Contents

Abstract

Although there are many machine learning programs that can acquire new problem solving strategies, we do not know exactly how their processes will manifest themselves in human behavior, if at all. In order to find out, a line-by-line protocol analysis was conducted of a subject discovering problem solving strategies. A model was developed that could explain 96% of the lines in the protocol. On this analysis, the subject's learning was confined to 11 rule acquisition events, wherein she temporarily abandoned her normal problem solving and focused on improving her strategic knowledge. Further analysis showed that: (1) Not all rule acquisition events are triggered by impasses. (2) Rules are acquired gradually, both because of competition between new and old rules, and because of the subject's apparently deliberate policy of gradual generalization. (3) This subject took a scientific approach to strategy discovery, even planning and conducting small experiments.

1. Introduction

A decade ago, cognitive science "solved" human problem solving. Explicit computational models were developed that could solve difficult puzzles, and convincing evidence was found for their psychological reality. The most convincing evidence came from line-by-line analysis of concurrent protocols (e.g., Newell & Simon, 1972; Ohlsson, 1980; Karat, 1982; see VanLehn, 1989a, for a review). Of course there was controversy about important details (e.g., Was the programming formalism used in the simulations a good model of the cognitive architecture?[1]). Yet there was little doubt that people solved novel problems using means-ends analysis, forward search, abstraction planning, and other weak methods.

It was soon found that experts often used distinctly different problem solving strategies than novices, so strategy acquisition became an important topic. In the late 1970's, the first computational models of strategy acquisition began to appear, and the field of machine learning was born. Nowadays, many models exist and more are invented each year (130 papers were presented at the 1989 machine learning workshop).

However, strategy acquisition is not "solved" in the way that problem solving is, because the evidence connecting models to human data is weak. Most of the existing work reduces human data to a sequence of strategies, then demonstrates that the model can make these transitions too (Neves, 1978; Anzai & Simon, 1979; Lewis, 1981; Anderson, Greeno, Kline, Neves, 1981; Larkin, 1981; Anzai, 1987; Neches, 1987; Langley, 1987; Ohlsson, 1987; Wallace, Klahr & Bluff, 1987; VanLehn, 1990). A few studies have reduced the data to a "protocol abstract," which is a sequence of episodes lasting a few minutes each (Anderson, Farrell, & Saurers, 1984). In all this research, the data are so reduced that they place just two constraints on the strategy acquisition process. First, the machine's learning process has the same input and output as the human's learning process. That is, they make the same strategy transitions. Second, the machine uses roughly the same information as the subject. For instance, if the subject refers to worked example solutions, then the machine would do so as well.

These two empirical constraints leave the exact nature of the acquisition process vastly underdetermined. As just one illustration of this underdetermination, consider the classic work by Anzai and Simon (1979). Using a concurrent protocol, Anzai and Simon showed that a subject solving the Tower of Hanoi puzzle acquired four strategies consecutively while receiving no feedback from the experimenter. They developed a computer model that could make the same strategy transitions without

feedback. On this basis, they argued for the model's psychological plausibility. In 1979, there were no competing models, but now, three machine learning models have made some or all of the strategy transitions found by Anzai and Simon (Langley, 1985; Anderson, 1989; Ruiz & Newell, 1989). The new models are not just minor variants of the Anzai and Simon model. The Anzai and Simon model monitored working memory in order to detect bad patterns of moves. The Langley (1985) model analyzed search trees and constructed heuristics for avoiding bad paths using a concept formation technique. The Anderson (1989) model uses a sequential connectionist network. The Ruiz and Newell (1989) model is based on Soar, with its chunking mechanism for learning. Although all these models made the observed transitions without feedback, they utilized extremely different processes. Moreover, there are many other programs in the machine learning literature that could also make of the observed transitions, even though the authors chose other tasks than the Tower of Hanoi for demonstrating them (e.g., Mitchell, Utgoff & Banerji, 1983; Minton et al., 1989). In short, the early belief that we understood how subjects acquired new strategies in the Tower of Hanoi and other classic tasks has been undermined by the explosion of machine learning models.

There is an even more important area of uncertainty in the current understanding of strategy acquisition. Machine learning programs often take many steps and employ a variety of mechanisms. We do not know how these steps and mechanisms map onto human behavior. As a specific example of our ignorance, consider again the Anzai and Simon (1979) model for strategy acquisition. It looks for patterns in its recent reasoning, which is recorded in its working memory. If it sees a pattern of operator applications that lead to a bad state, it builds a rule that prevents that sequence of operators from being used in the future. Conversely, if it detects operator applications leading to good states, it builds rules that try to make them occur again. If it finds that certain patterns of operator applications that are repeated frequently, it converts them into a single macro-operator or chunk. Although the model is clear and simple, it is not clear how its processes map onto human behavior. When a bad-state pattern is detected; would the subject pause for a few seconds then mumble "I'd better not do that again," or would the process take place unconsciously, or at least with no visible pauses or comments by the subject?

The question addressed by this research is simply this. If we could isolate some cases of a strategy acquisition process in action, what would the person's verbal protocol look like? In particular, would the subject interrupt her normal activities in order to run the strategy acquisition process, or would

the strategy acquisition process run automatically in the background, and thus generate no verbal signs of its activities? Suppose further that we can find *all* occasions when a person's strategy changes. Would all of them manifest themselves as interruptions of normal processing, or only some? If some strategic changes are evident in the protocol and some are not, is there any characteristic that distinguishes the two cases?

This paper presents a line-by-line analysis of the protocol of a subject acq 'ring a sequence of strategies. The analysis was not conducted as a test of any particular machine learning model. The primary goal was simply to understand the mapping between machine learning models and protocols of human behavior. This work has the same goals (and the same attention to detail!) as the protocol analyses of Newell and Simon (1972). By 1972, their group had already published a computational treatment of problem solving (Ernst & Newell, 1969), so it appears that the primary goal of the 1972 book was to explicate the mapping between human behavior and their models. Similarly, the primary goal of the present treatment is also to connect models and protocol data, although in this case the models are drawn from the machine learning literature. The secondary goals of Newell and Simon (1972) appear to have been to the augmentation of psychological theory and computer science technology. These are the secondary goals here, too. In this case, two new machine learning techniques are suggested (scientific strategy acquisition, and a combination of perturbation-based and explanation-based learning) and three psychological issues are addressed (Is all skill acquisition driven by impasses? Why do discovery learners forget what they discover? What makes good students good?).

2. The plan of the analysis

This study is a reanalysis of the classic protocol of Anzai and Simon (1979) wherein the subject invents several solution strategies for the five-disk Tower of Hanoi over the course of 90 minutes. During this time she receives no instruction. This protocol was selected for study because it is known to encompass significant learning and because the subject gave an unusually clear protocol.

My analysis is a refinement of Anzai and Simon's original analysis. Anzai and Simon uncovered the major strategies that the subject acquired and postulated learning mechanisms sufficient to acquire those strategies. They did not attempt a line-by-line comparison of the protocol and the behavior of their model.

Given that the goal of this analysis is to find out how strategy acquisition processes manifest themselves in protocol data, there are two subgoals: Find the lines in the protocol where the processes are active, and see what happens there. The plan for achieving the first subgoal has three sub-subgoals:

1. *Find out what rules the subject has.* Following Anzai and Simon (1979), we assume a rule-based representation of strategies. This implies that the job of precisely specifing the subject's strategies amounts to developing a set of rules that explains the subject's actions and utterances. Anzai and Simon did most of the work here, so their analysis is presented first, in section 3.1. However, a few more rules are needed in order to explain a subtle pattern in the subject's verbalizations. The pattern and rules are presented in section 3.2.

2. *Find out where each rule fires.* (We also need to find missed opportunities -- places where the rule could have fired but did not.) This step of the analysis is purely mechanical. A computer was used to enumerate all possible rule firing sequences and select the one that maximizes the fit with the subject's utterances. Section 3.3 presents the best fitting rule firing sequence.

3. *Infer the location of the strategy acquisition processes.* If a rule missed all opportunities to fire prior to line N of the protocol, and it began to fire regularly starting with line N, then the rule was probably acquired in the vicinity of line N. Section 3.4 presents an analysis based on a narrow definition of "vicinity." Section 4 uses a broader and more successful definition.

These three sub-subgoals locate the acquisition processes, thus achieving the first subgoal of the analysis.

The second subgoal is to characterize the subject's overt behavior during the execution of a strategy acquisition process. It could be that the protocol looks exactly like it does at any other time. This would suggest that strategy acquisition processes do not require the subject's attention, and thus can run automatically in the background while the subject pursues her problem solving goals, as suggested by Schoenfeld, Smith and Arcavi (in press) and others. Alternatively, it could be that the subject interrupts her normal problem solving, switches her attention to the problem of modifying her strategy, and begins to mutter goals appropriate to that task instead of her normal disk-moving goals.

This kind of behavior would suggest that strategy acquisition processes are a form of problem solving that is directed at the meta-problem of improving a task's solution method, as suggested by Karmiloff-Smith and Inhelder (1974) and others. In short, the subject's verbal behavior in the vicinity of a rule's first firing should tell us something about how strategy acquisition processes surface in human behavior. Section 4 presents this part of the analysis.

3. Locating the protocol lines where strategies are acquired

3.1. The Anzai and Simon analysis

Throughout this paper, most of the nomenclature, line numbers and episodes of Anzai and Simon are retained. In their system, the goal of the Tower of Hanoi puzzle is to move five disks from the initial peg, peg A, to a final peg, peg C. There is one other peg, peg B, that is used to hold disks temporarily. The disks vary in size, and larger disks are given larger numbers. Disk 1 is smallest and disk 5 is largest. When stacked in numerical order, the disks form a pyramidal object, which is conventionally identified as the N-high pyramid, where N is the size of the largest (i.e., the bottom) disk. A single move consists of removing a disk that is exposed on the top of a peg and placing it on a different peg. However, a larger disk cannot be placed on a smaller disk.

Anzai and Simon divide the protocol into four episodes. Appendix 1 of this article presents the protocol, arranged in an unusual way that will be explained later. During the first episode (lines 1 to 24), the subject attempts to solve the five-disk puzzle and gives up after 11 moves. The subject infers, correctly, that her failure was caused by making the wrong initial move. During the second episode (lines 25 to 74), she deliberately makes the opposite initial move, and eventually succeeds in solving the five-disk puzzle. At the beginning of the third episode (lines 75 to 107), the subject solves successively larger versions of the puzzle, starting with the trivial 1-disk puzzle, and proceeding to the 2-disk puzzle, the 3-disk puzzle, and the 4-disk puzzle. The third episode ends with a correct solution of the original 5-disk puzzle (lines 108 to 162). At this point, the subject seems happy with her solution strategy, and only attempts the puzzle again because the experimenter asks her to. The fourth episode (lines 163-224) consists of her final, correct solution to the puzzle.

All the moves made by this subject, except for the first move of the whole protocol, were optimal. Evidence for learning must therefore come from the pauses and the verbalizations of the subject, and not from the correctness or incorrectness of the moves. More importantly, the fact that virtually all her

move selections were optimal suggest that she changed strategies in order to improve the efficiency or elegance of her problem solving, since she was already obtaining a correct solution. This is an important property of this particular protocol that crops up repeatedly in the analyses. One could classify this protocol as an instance of discovery learning (because no feedback was given) that is driven a desire to improve the elegance or operationality of a theory, rather than its empirical coverage.

In seeking to improve her solution strategy, the subject was just following the instructions given to her by the experimenter. The last two lines of the instructions are: "If you think that your solving process would not lead to a good solution procedure, you may give up that process and start from the initial situation. I hope that you can find a good solution procedure for the problem." (Anzai & Simon, 1979, pg. 125). As Anzai and Simon point out, the instructions encouraged the subject not merely to find a solution to the puzzle, but to find a good solution procedure.

On Anzai and Simon's analysis, the subject uses the following four solution strategies in succession:

1. *Selective search:* The subject only considers moving disks that are free to move in the current state. However, she uses two heuristics: (1) do not move the same disk on consecutive moves, and (2) do not move the smallest disk back to the peg it was on just before it was moved to its current peg. The later heuristic requires recalling where the smallest disk was two moves earlier.

2. *Goal-peg:* This seems to be a mixture of the preceding strategy and the one that follows. Anzai and Simon's description is: "In the second episode, unlike the first, the subject guided herself explicitly by mentioning intermediate goals: to move Disk 4 to Peg B (line 34), to move Disk 5 to Peg C (line 48), to move Disk 4 to Peg C (line 59), and to move Disk 3 to Peg C (line 63). She summarized this strategy of moving first the largest and then the successively smaller disks to C in lines 72-74. We refer to this as the *goal-peg strategy.*"

3. *Disk subgoaling:* This is a classic means-ends analysis strategy for solving the puzzle. To plan a move, the subject focuses on the largest disk that is not yet on peg C. She determines which disks are blocking the movement of that disk to peg C. She focuses on the largest blocking disk, and decides which peg it would have to be on in order to allow

the movement of the original disk to peg C. Putting this blocking disk on that peg now becomes a new goal, and the strategy recurses. For instance, in lines 82-84, the subject says "Oh yes, 3 will have to go to C first; for that, 2 will have to go to B; for that, um..., 1 will go to C."

4. *Pyramid subgoaling:* This is also a recursive solution strategy, but the subject thinks in terms of moving pyramids instead of disks. For instance, in lines 210-212, she says "Next, if the three at A go to C, I will be done. So first, the top two disks will be moved to B. For that, 1 goes from A to C."

The numbering above also reflects the rough location of the strategies in the protocol: selective search is used in episode 1, goal-peg in episode 2, etc. However, Anzai and Simon do not say exactly where in the protocol the transitions between these strategies takes place. In order to do that, a finer-grained analysis is required. The next sections develop such an analysis.

3.2. The 4k+1 pattern

Because the subject solves the five-disk puzzle optimally three times, in episodes 2, 3 and 4, she makes exactly the same sequence of moves three times. In order to see how the subject's strategy changes, one can align the protocols from episodes 2, 3 and 4, as shown in table 1. The first column is the move number. The second column will be explained in a moment. The remaining three columns summarize the subject's utterances in episodes 2, 3 and 4 respectively. The alphanumeric symbols stand for goals mentioned by the subject. Thus, 3B means the subject said something like "Next, get 3 to B." The ellipses in the table stand for extended non-goal comments mixed with long pauses. An expanded version of this table is presented in an appendix, with the actual text in place of the abbreviations used in table 1.

Place table 1 about here.

Before beginning the analysis of table 1, it is worth emphasizing the nature of the data to be explained. The subject's learning does not affect her solution path at all. It only shows up in her verbal utterances. Thus, the changes of the goal statements shown in table 1 are the crucial data to explain.

Table 1 demonstrates that there is a pattern that cuts across all three episodes. Although most of the moves are brief remarks indicating the move itself (e.g., "1 will go from A to C"), some of the moves

are marked by extra talk on the subject's part. There is a simple pattern. With one exception, all the extra talk occurs on moves 1, 5, 9... 4k+1. This pattern holds across all three episodes. The probability of this pattern occurring by chance is less than 0.001, by Chi-square test.

This pattern is not consistent with the strategies of Anzai and Simon.[2] On their analysis, the subject uses the selective search strategy during episode 2 and the two subgoaling strategies during episodes 3 and 4. If the subject is using the selective search strategy during episode 2 (i.e., do not move the same disk twice in a row; do not move disk 1 back the the peg it came from), then all of the moves except the initial move should be brief, perfunctory comments because the selective search strategy involves no subgoaling or planning. However, the subject seems to struggle at each of the moves 4k+1 in episode 2, contrary to the prediction of the selective search strategy.

Likewise, if the subject is using either the disk subgoaling strategy or the pyramid subgoaling strategy during episodes 3 and 4, then she should mention more goals than she does. Column 2 of table 1 indicates the goals that the subgoaling strategies would produce. Column 2 shows a goal explicitly only when it is created; ditto marks indicate goals stored in working memory. If the subject followed the convention of always mentioning a goal, even if it is stored in memory, then she would mention many more goals than she does. If she followed the convention of only mentioning the new goals, then she should mention goals at moves 13, 21 and 29 that she does not. Also, she would not mention goals at moves 5 and 9 that she does in fact mention. So there seems to be no simple explanation for her utterances based on the assumption that she is following the subgoaling strategies throughout episodes 3 and 4.

More to the point, there is an obvious pattern that cuts across all the episodes -- the extra talk only occurs on moves 4k+1 -- and the Anzai and Simon analysis does not explain it. The simplest hypothesis is that the stability of this pattern is due to rules that are stable across all the episodes. It is easy to construct such a set of rules. The following is one possibility:

- It is inappropriate to move the same disk on consecutive moves. That is, if you just moved, say disk 2 to peg B, and you intend to continue forward along this search path rather than back up, then do not move disk 2 off peg B onto another peg; move some other disk instead.

- If there is only one appropriate action, then do it.

- If there are multiple appropriate actions, but one of them is to put disk 1 on top of disk 2, thus forming a 2-high pyramid, then do that action.

A well-known property of the Tower of Hanoi is that the first two rules uniquely determine the choice of action on all the odd numbered moves. Because third rule takes care of moves numbered 4k+3, the three rules together uniquely determine the choice of action on all the moves *except* moves 4k+1. At each of moves 4k+1, there are two appropriate actions and neither involves forming a 2-high pyramid, so these rules underdetermine the choice of action at moves 4k+1. These three rules are quite simple to apply, as they require neither subgoaling nor heavy demands on working memory. If the subject had these rules, then the only time she would need to reason deeply is when these rules underdetermine the choice of action. Thus, she would tend to do all her talking at moves 4k+1, which is exactly what we observed in table 1.

Although I find this particular set of rules quite plausible, they are not the only possible explanation for the 4k+1 pattern. For instance, it could be that the subject has the first two rules plus a chunk, routine or plan consisting of three steps: take the forced move, make a 2-high pyramid, and take the forced move. It is not possible to tell which account is correct. Both the three rules and the chunk are 100% accurate in predicting the subject's actions at moves 4k+2, 4k+3 and 4k. The verbal data at those moves is too weak too help, for they consist merely of perfunctory announcements of actions. Fortunately, this weakness in the data does not harm the main argument of this paper, which is aimed at explaining the *changes* in the subject's rules. Since the subject's strategy for making these moves does not change during the course of the protocol, it does not matter whether it is composed of the three rules cited above or some other kind of knowledge.

In addition to the three rules above (or their equivalent), I assume that the subject uses the following rule throughout the protocol:

- The top level goals are to first get disk 5 to peg C, then get disk 4 to peg C, then disk 3, disk 2 and finally disk 1.

The use of this rule is supported by ample verbal evidence (lines 13, 48, 59, 63, 72, 78, 99, 110, 129, 140, 153, 165, and 194). It is possible that the subject inferred it from the puzzle's instructions.

3.3. The major moves

The overall objective is to find the protocol lines where the subject changes strategy. The 4k+1 pattern indicated that there is a stable set of rules (or plans, etc.) that handle many of the moves in the protocol. To be exact, there are 130 moves in the protocol, and the rules assumed in the preceding section explain the subject's actions and utterances on 95 of them. The remaining 35 moves include moves 4k+1 from each of episodes 2, 3 and 4, and some additional moves from episodes 1 and ⁊ Because the 95 moves are handled by a strategy that does not change, any strategy changes that exist in this protocol will show up in among these 35 moves. For handy reference, let us call these 35 moves the *major moves*. The next step is to find out which of these major moves are the locations of the subject's strategy transitions.

Table 2 shows the ten rules on which the analysis is based. The first four rules are the ones presented in the preceding section. The other six rules were constructed post-hoc for explaining the major moves. Three of them, labeled "initially present rules," seem to have been inferred by the subject from the instructions or learned very early in the protocol, because there is evidence that they were used on the first possible occasion where they could be used. Rule SSS, which is the selective search strategy discussed by Anzai and Simon (1979), could have been inferred in the ways they describe. Rule 1blk is just an instantiation of a common sense planning rule: If you want to move something, and something else is in your way, then move the blocking thing out of the way. The origins of rule 2blk are not clear. It can be deduced by two applications of rule 1blk or induced by the mechanisms of Anzai and Simon (1979). The protocol in the neighborhood of the rule's first use (lines 30 to 34) shows a good deal of hesitation and search, but the subject's comments are not sufficiently clear to indicate what acquisition mechanism was responsible for the rule's formation, if indeed it was formed at that move. Interestingly, the concept of a 2-high pyramid plays a central role in both rule 2blk and rule put-1-on-2. This concept may have been made salient by something in the instructions, which would explain why both rules appear so early in the protocol. The actual instructions given to the subject (see Anzai & Simon, 1979, pg. 125) do not mention it, however.

The last three rules in table 2 appear to have been acquired during the course of the protocol. As their acquisition is the topic of several later sections, no justification will be offered for them here. (These renditions of the three learned rules are not completely accurate. They will be refined later.)

Place table 2 about here.

Place table 3 about here.

The analysis of the major moves is shown in table 3. Each row is a major move. The first column numbers the major moves. (This numbering has nothing to do with the numbering of table 1.) Horizontal lines in the table indicate places where the subject reset the puzzle to an initial state. Episode 1 corresponds to major moves 1 through 3, episode 2 to major moves 4 through 11, episode 3 to major moves 12 through 27, and episode 4 to major moves 28 through 35.

The second, third and fourth columns summarize the protocol. The second column abbreviates the puzzle's state just prior to the move. The notation "125,34,_" means that disks 1, 2 and 5 are on peg A, disks 3 and 4 are on peg B, and peg C is empty. The third column abbreviates what the subject said while making the move. The notation "2B, 1A" means that the subject announced a goal of moving disk 2 to peg B, then announced a movement of disk 1 to peg A. The notation "4pC" indicates a goal of moving a pyramidal group of four disks. Sometimes the subject announces a series of goals, pauses, and announces a different series of goals. This behavior is indicated by placing two rows in the table, one for each series of goals, and placing ditto marks in the first two cells of the second row (see major moves 18 and 21). The fourth column is a rough indication of the subject's non-goal utterances. The numbers count the words and pauses, exclusive of goal and action announcements. The minus sign indicates that the subject's remarks are negative (e.g., "Um... it's hard, isn't it?") or just showed long pauses. For instance, the remark, "For that, um...., this time, again..., as this time 4 will have to go to B..." would receive 12- as its code because there are 3 pauses and 9 words, exclusive of those announcing the 4B goal. If the subject's remarks were generally positive, the word/pause count is followed by a plus sign. If the subject mentioned only her action and goals, perhaps with a small pause or connecting phrase, then the cell is left blank.

The rightmost six columns of the table indicate which rules were fired on which moves. These columns are coordinated with table 1. The four stable rules, which appear at the top of table 1, are not shown in table 3. Only the six rules whose firing pattern changed are shown. Their columns are labeled with the abbreviations from table 1. The cells in these columns indicate rule firings. Given a cell in a certain column and a certain row, a blank in the cell indicates that that the column's rule was not applicable during the row's move, a "0" indicates that the rule was applicable but did not fire, and a "1" indicates that the rule was applicable and fired.

In order to check the major move analysis, a series of calculations were performed.[3] Each row was checked separately. The first step was to calculate all possible derivations, where a derivation is a sequence of rule firings that eventuates in selecting an action. Some of these derivations selected disk-moving actions that did not correspond to the subject's actions, so we can infer that this sequence of rule firings is not what the subject did. For instance, the disk subgoaling strategy would have selected 1C as the action for major move 1, and that is not the action the subject chose. Of the remaining derivations, some predicted different goal utterances that those made by the subject, so we can eliminate these derivations as well. For instance, at major move 21, the disk subgoaling strategy predicts the goal sequence 5C, 4B, 2C, 1A. This is not what the subject said. After derivations inconsistent with the protocol had been removed, there would be either 0, 1 or more than one derivation left. The more-than-one case never occurred, as it turned out. Most frequently, there was only one derivation left, which allows us to assume with confidence that this particular sequence of rule firings is indeed the one employed by the subject for this major move. The rules involved get a 1 placed in their cell in the table 3. Rules involved in the other derivations get a 0 placed in their cells. The other cells on this row are left blank.

The remaining possibility is that no derivations are left after filtering. This means that all the derivations calculated for a major move were inconsistent with the subject's actions and utterances. This occurred five times. However, in all five cases, there was a derivation that came close to matching the utterances, so it was used in formulating table 3. It is worth a moment to discuss these five cases. In the first two cases (major moves 30 and 33), the subject did not mention a goal that she should have. This is probably just a lack of attention or a slip of some kind. In the another case (major move 10), she mentions a top level goal (moving disk 3 to peg C) for reasons unknown, and similarly at major move 2, she mentions a goal (moving disk 2 to peg B) for reasons unknown. In the last case (major move 3), she mentions a top level goal (moving disk 5 to peg C) that the analysis cannot explain. However, learning seems to be going on at this move (according to the analysis presented later), and mentioning that goal seems to be a part of the reasoning involved in formulating the new rule.

The fit between this analysis and the protocol rivals some of the best in the literature. For instance, in the famous Newell and Simon (1972) analysis of S3 on a cryptarithmetic problem, there were 267 nodes in the problem behavior graph derived from the subject's protocol (Newell & Simon, 1972, pg. 202-203). Of these, their simulation generated 237 nodes, but it also generated 23 nodes that

were not in the problem behavior graph. Newell and Simon evaluated this as a fit of (237-23)/267 or 80%. In our Tower of Hanoi analysis, nodes in the problem behavior graph are equivalent to the subject's overt moves and goal utterances. The subject made 130 overt actions and 41 goal utterances, for a total of 171 "nodes." The major move analysis and the analysis of the 4k+1 pattern explain all but one of the overt actions. (The action they do not explain is the last move of episode 1, where the subject changes her mind and undoes her previous move.) The major move analysis explains all the goal utterances in table 3 except for the extra utterances at moves 2, 3 and 5. There is one goal utterance at a non-major move (the 22nd move of episode 2 -- see table 1) that is not explained. Thus the analysis explains 41-4 goal utterances, and (130-1)+(41-4)=166 of the 171 nodes. However, the analysis predicts that the subject will utter two goal utterances that she does not (major moves 30 and 33). Using Newell and Simon's calculation, the analysis fits (166-2)/171 or 96% of the protocol. Although this method for calculating fit is quite ad hoc,[4] it gives one a rough feel for how well the analysis fits the protocol. In a word, the fit is excellent.

The analysis presented so far indicates that the subject acquired three rules. Moreover, it has located the lines in the protocol where the first use of the rule occurred. The next step in the analysis is to examine the protocol at those locations, since they are likely to be where the subject's strategy acquisition processes are active.

3.4. What happens at the initial firings of the acquired rules?

Rule 4B is first used at major move 5, rule Dsk at major move 14 and rule Pyr at major move 30. Table 4 shows the relevant sections of the protocol. The puzzle states appear in the left column.

Place table 4 about here.

These segments of protocol are rather unilluminating. Although there is clear evidence for each rules firing (4B at line 34, Dsk at lines 82-84, and Pyr at line 179), and there is some extra talk surrounding the firings of rule 4B and Dsk, it is not at all clear what the extra talk means.

Apparently, the initial firings of the rules are not the only places where strategy acquisition is taking place. In all three cases, it appears that some learning took place *before* the initial firing. Apparently, when we use the heuristic of looking in the vicinity of the initial firing, we should use a wider definition of "vicinity" than just the protocol at that single move. The next section does exactly that.

4. Further analysis of the protocol

In order to find the places where strategy acquisition occurs, a broader kind of analysis is used. In order to constrain it, the new analysis must use the actual content of the non-goal talk, which has so far been ignored (Table 3 is based on the actions and goal utterances alone.[5]) In order to interpret the non-goal talk, a overarching pattern in the data needs to be pointed out. The next section discusses it, then subsequent subsections discuss the acquisition of each rule in turn.

4.1. Scientific strategy discovery

The subject's overall approach to strategy acquisition method to be the classic scientific method of hypothesis formation and experimentation. Although this may be just an idiosyncrasy of this subject, it is important to explicate it in order to interpret this protocol. Also, this aspect of her behavior is just plain interesting, given our current ignorance about the cognitive processes involved in discovery.

The subject's investigation begins just after she successfully solves the puzzle for the first time. As she moves the last disk to peg C, she says (lines 70 and 71), "All right, I made it. I wonder if I've found something new." She then generates some observations about her recent solution (line 72): "I don't know for sure, and little ones will have to go on top of big ones... big ones can't go on top of little ones, so first, bit by bit, C will be used often before 5 gets there. And then, if 5 went to C, next I have to think of it as 4 to go to C..." Although none of her observations seem to lead anywhere, in a sense, she is acting like a scientist whose first act is to marshal the known evidence about a phenomenon in order to see if they suggest a hypothesis.

As mentioned earlier, the instructions to the subject encourage finding a *good* solution procedure. But what does "good" mean to the subject? Anzai and Simon (1979) hypothesized that the subject's motive was to find an *efficient* solution path. Since 1979, a great many machine learning programs have followed Anzai and Simon's lead and used efficiency as their major motivation for strategy acquisition. However, a close inspection of this subject's verbal behavior suggests that her motive may have been to *understand* the solution path rather than to increase her efficiency. She does not say, "I wonder if there's a better way to do this," and start to propose shortcuts of various kinds. Since she never had to backup during episode 2, she might think that she had already found the optimal solution path (which she had) and thus believe that no great increases in efficiency were possible. Instead, she seems to sense that the solution path has an underlying mathematical structure (which it does) and sets

about to discover it.[6]

Because the subject's initial marshaling of the evidence fails to generate a worthwhile hypothesis, she takes the next step in the scientific method: she tries to reproduce the phenomenon in a simpler setting (cf. Kulkarni & Simon, 1988). In fact, she plans a whole series of experiments with increasing complexity. She starts with the simplest puzzle possible: one disk on peg A. She solves the puzzle (line 76), but seems to learn nothing from it. She sets up the two-disk puzzle and quickly solves it (line 77). Then she applies a crucial scientific heuristic, and reflects on the results of her experiment (line 78): "That... that anyway, 2 will have to go to the bottom of C, naturally I thought of 1 going to B." The word "bottom" indicates that she sees why she should first concentrate on getting disk 2 to peg C -- once there, it will not have to be moved again. The word "naturally" indicates that she sees an intimate connection between the goal of getting 2 to C and the movement of 1 to B. This seems to be the beginning of a hypothesis. Note that she already "knew" these rules in some form, because she just used them to solve the puzzle. What she seems to be doing here is picking out these pieces of knowledge from among the many she has used, and considering whether they might be the key to understanding the puzzle. Granted, this interpretation is going far beyond the verbal evidence, but let us pursue it a little further. Later verbal reports provide more support for this interpretation.

The next step in her research plan is to investigate the three-disk puzzle, so she sets it up and says: "So if there were three..., yes, yes, now it gets difficult. Yes, it's not that easy.... This time, 1 will...." This puzzle should not be difficult for the subject, since she has two rules that suffice to calculate an initial move (rules SSS and 2blk). Indeed, just a few lines earlier, she easily solved the isomorphic problem of calculating a move for the state [123,_,45]. Apparently, she is deliberately ignoring these rules in order to concentrate on her new idea.

At the next line, 80, she finally sees the parallels between this situation, [123,_,_], and the preceding one, [12,_,_]. She says, "Oh, yes, 3 will have to go to C first," presumably because disk 3 will be on the bottom of peg C. On the next line she says, "For that, 2 will have to go to B." This appears to be a generalization of her earlier insight (line 78) about why one should move disk 1 to peg B, given that one wants disk 2 on peg C. At this point she has found the basic ingredients of the disk subgoaling strategy.

Her crucial step was to avoid two perfectly good rules, SSS and 2blk, in order to find another way

to solve the puzzle. It is likely that she *deliberately* avoids these rules, rather than failing to retrieve them. The rules were used consistently throughout the whole of episode 1 and 2, so they should be familiar (strong or multiply encoded, depending one's theory of memory) and easy to retrieve. Indeed, the rules appear to be so easy to retrieve that she has trouble implementing her policy of avoiding them. At several points she slips and starts use 2blk. On two occasions, major moves 18 and 21, she catches her slip, recovers and applies her new strategy. On several other occasions (major moves 23, 25, 29 and 31), she fails to detect her slip, and solves the move with the older rule. Thus, it appears that the subject is trying to avoid the old rules but occasionally lapses back into using them anyway.

In general terms, this subject's behavior is similar to classic cases of discovery behavior (e.g., Inhelder & Piaget, 1958; Siegler, 1983; Shrager & Klahr, 1986). For instance, Karmiloff-Smith and Inhelder (1974) observed young children discovering how to balance blocks on a narrow fulcrum. After a few initial attempts at balancing blocks (equivalent to episode 2 in this protocol), the subjects would more-or-less ignore the balancing goal and undertake a detailed exploration of the properties of blocks (equivalent to the early parts of episode 3). Eventually, they would notice (cf. Ruiz and Newell, 1989) that some of the blocks could be balanced by placing their geometric centers directly over the fulcrum (equivalent to the discovery at line 78). This new strategy is gradually modified as its inadequacies are discovered (equivalent to the generalization of the disk subgoaling strategy). Some of the older subjects planned a series of block balancing "experiments" so as to minimize the number of block properties that changed in successive balancing attempts (equivalent to the subject's design of a succession of puzzles of gradually increasing size).

Although the details of this analysis are easy to dispute, two facts are plain: (1) With no prompting from the experimenter, the subject solved puzzles of increasing size. (2) She had difficulties which she should not have had if she was using all the rules that she used during her earlier solution of the 5-disk puzzle. These facts suggest that she was applying some kind of deliberate strategy acquisition method, and that method is strikingly similar to those used by scientists and mathematicians.

This assumption about the subject's overall approach to learning plays an important role in interpreting the subject's remarks while learning individual rules. The next three sections present analyses of the three rules learned by the subject.

4.2. Acquiring rule 4B

Rule 4B is first used at major move 5, but the subject seems to learn it at earlier, at major move 3. The protocol there is:

[45,123,_]	11.	So then, 4 will go from A to C.
[5,123,4]	12.	And then..., um..., oh..., um...,
	13.	I should have placed 5 on C.

At line 11, the subject has gotten the puzzle into the state [45,123,_]. (In this and subsequent excerpts from the protocol, the puzzle's state appears in the left column.) During the long pause at line 12, the subject seems to realize that placing disk 5 on peg C means that disk 4 has to go to peg B first, and this causes her to form rule 4B.[7]

The acquisition of rule 4B seems to be triggered by an *impasse*. Although term "impasse" can be defined precisely only in the context of precisely defined cognitive architecture, its approximate meaning is that the architecture cannot decide what to do next given the knowledge and the situation that are its current focus of attention (Brown & VanLehn, 1980; Laird, Newell, & Rosenbloom, 1987). In order to avoid technicalities at this stage in the analysis, an atheoretical criterion for "impasse" will be used. Let us assume that an impasse has occurred wherever the subject makes a negative statement about her progress (e.g., "Oh, this won't do...", "Oh no! If I do it this way, it won't work.", "No not 2, but I placed 1 from B to C...right?") or the subject pauses twice (indicated by ellipses in the protocol) with no intervening goal utterance or overt moves. In table 3, all the major moves that have minus signs in the "Extra" column have an impasse, according to this criterion.

To return to the discussion of rule 4B, it appears that the subject reaches an impasse just after moving from state [45,123,_] to [5,123,4]. One explanation of this impasse is that she focuses on the goal of putting disk 5 on peg C in the state [5,123,4] where only disk 4 blocks the move. This triggers her common sense rule about moving an object out of the way (rule 1blk -- see table 2), so she formulates the subgoal of moving disk 4 to peg B. This subgoal cannot be immediately achieved, because peg B is occupied by smaller disks than 4. Thus, she is at an impasse.

The subject's comment, "I should have placed 5 on C," sheds no light on the type of reasoning she uses during this acquisition event. It would be computationally simple for her to deduce rule 4B at this point. For instance, she could have reasoned:

- Because getting disk 4 out of the way on B will always be a subgoal of moving disk 5 from A to C (by rule 1blk),

- and because moving disk 5 to C is always the first top level goal to be achieved in the five-disk puzzle (by the rule that states that the top level goals of the puzzle are to get 5 to C, then 4 to C, etc.),

- it follows that moving disk 4 to B is a prerequisite to achievement of any of the top level goals of the five-disk puzzle.

Whether or not the subject makes this deduction is unclear from the protocol. However, it is fairly clear that she does it in response to an impasse.

There are several machine learning models that learn from failures using deductive reasoning (e.g., Prodigy; Minton et al. 1989). Most of them do not learn at an impasse (the point of failure) but instead wait until a complete, successful solution is found before going back to find the failures and learn from them. Waiting has the advantage that one can intelligently choose which failures to learn from and thus save learning effort. The disadvantage of waiting is that it requires storing the whole search tree in memory. This subject does not wait until she has found a solution to the puzzle. Indeed, at major move 3 she is on a failing branch of the search tree. But she apparently learns a rule here anyway. So it is clear that this subject is learning from impasses, rather than using the more common wait-and-see approach to failure-driven learning.

The next possible opportunity for rule 4B to fire is major move 4. The subject has just reset the puzzle to its initial state. However, the subject does not say something like "Now I need to get 4 to B, so I'll...," which would indicate a firing of rule 4B. Instead, the subject says: "Since 1 is the only disk I can move, and last I moved it to B, I'll put it on C this time." The subject recalls that she started out her disastrous first attempt at the puzzle by moving 1 to B, so this time she moves it to C. Rule 4B is ignored.

At the next major move, the subject clearly indicates a firing of rule 4B when she says, "Because....I want 4 on B,...." (line 34). Why did the subject use rule 4B here and not on the preceding major move? As table 3 shows, rule 4B is only used on moves where the puzzle is in the state [45,12,3]. It is not used at the only other place where it could be used, where the puzzle is in the state [12345,_,_]. This suggests that the version of rule 4B that the subject actually has is much more specific than the one in table 2. It might include conditions that cause it to fire only when, say, disk 4 is exposed on top of a peg. Such overspecificity is the hallmark of certain inductive learning algorithms,

such as the famous wholist procedure of Bruner, Goodnow and Austin (1956). This type of learning, and not the deduction mentioned earlier, might be responsible for the acquisition of rule 4B. On the other hand, it may be that rule 4B is perfectly general, but the visual cues of the state [45,12,3] facilitate retrieval where those of state [12345,_,_] do not. The data are not sufficiently rich to allow us to choose among these explanations.

In summary, the short career of rule 4B starts with some kind of impasse-driven learning around major move 3. This produces a rule that is probably more specific than the one shown in the table. Overspecificity, either in the rule's retrieval cues or its conditions (if there is in fact a difference between these) cause the rule to fire on only half of the subsequent states where it could fire.

4.3. Acquiring the disk subgoaling strategy

The acquisition of the disk subgoaling strategy, rule Dsk, has a complicated explanation that provides a very accurate account of the verbal protocol. The explanation is based on the following assumptions: (1) Rules are have conditions that determine whether or not they will apply to the current state. (2) Conditions are logical expressions containing variables and constants. (2) The initial version of a rule's condition is as specific as possible. (3) The subject only generalizes the rule's condition when forced to. (4) The rule is only generalized enough to get it to apply to the given situation. (5) The generalization is limited to changing constants to variables and dropping conjuncts. These assumptions characterize the generalization methods of Sierra (VanLehn, 1987; VanLehn, 1990), early versions of ACT* (Anderson, 1983) and many machine learning programs.

Place table 5 about here.

Table 5 shows the sequence of rule versions that the subject seems to have held. On this analysis, the first rule, Dsk0, is learned at line 78. The subject has just finished solving the two-high puzzle. She reflects on her solution, saying

[_,_,12]   78.   That... that anyway, 2 will have to go to the bottom of C,
                           naturally I thought of 1 going to B.

The words "naturally" indicate that she thinks there is a connection between two moves, which is one indication that the rule is formed at this location (the major evidence is the success of the overall analysis of the acquisition of the disk subgoaling strategy). Dsk0 is a very specific rule. Indeed, it so specific that to call it a rule is a little presumptuous. Perhaps it should be called a case (Schank, 1982).

By the way, if this is indeed the beginning of the disk subgoaling strategy (and there is certainly room for disagreement, as the verbal evidence is quite weak), then we have a case of acquisition that is *not* triggered by an impasse. Not only does the subject show none of the official signs of an impasse (negative comments, long pauses), she has no reason to be stuck. She has just finished the 2-disk puzzle, and her next action should be to set up the 3-disk puzzle. There is nothing preventing her from doing that. But she pauses anyway, and reflects on her solution. It appears, therefore, that rule acquisition events can be triggered by successes as well as failures.

The next version of the disk subgoaling strategy, Dsk1s, seems to be learned at an impasse. After setting up the 3-disk puzzle, the subject says:

[123,_,_]   79.   So, if there were three..., yes, yes, now it gets difficult.
             80.   Yes, it's not that easy...
             81.   ...this time, 1 will...
             82.   Oh, yes, 3 will have to go to C first.
             83.   For that, 2 will have to go to B.
             84.   For that, um..., 1 will go to C.

Lines 79-81 are clear evidence for an impasse, since the subject makes both negative comments about her progress and pauses often without making goal utterances or overt moves. In order to have this impasse, the subject must be disregarding rules that she used earlier, in episode 2. In particular, she must not be using rules 2blk and SSS (see Table 2). This is consistent with her "scientific" approach to strategy formation.

The existence of an impasse here confirms our assumption that the rule (or case) formulated at line 78 is too specific to be applied to the state [123,_,_]. At the "Oh" of line 82, the subject appears to perform the minimal generalization necessary get the rule to apply to the state. The syntactically minimal generalization is to replace the constants, "disk 2" and "disk 1", with variables, "X" and "Y." This produces a new rule, Dsk1s.

The subject could have also turn the constants "peg A" and "peg C" into variables at the same time. This would not be a minimal generalization, because the constants happen to match the current goal, which is to get disk 3 from peg A to peg C. The main evidence that the subject chooses the minimal generalization instead of this one is that it explains impasses that occur later in the protocol. In particular, there is pause in the middle of line 84 that can be explained if she has Dsk1s, which explicitly mentions pegs A and C. This rule will not apply at line 84 where the goal is to move disk 2 to peg B. The overspecificity of Dsk1s explains the impasse at line 84.

If we assume that the subject again does minimal generalization at line 84, then she should get Dsk2s by replacing the constant "peg C" with a variable. Rule Dsk2s suffices for clearing blocking disks off peg A no matter how many of them there are. Thus, it should be able to handle major move 16. As predicted, the protocol shows no sign of an impasse:

[1234,_,_] 86. So, if there were four disks, this time, 3 will have to go to B, right?
            87. For that, 2 will have to stay at C, and then, for that, 1 will be at B.
            88. So 1 will go to B.

The first place that Dsk2s should fail to apply is major move 18. As predicted, the protocol shows signs of an impasse:

[_,123,4]   96. And then, again this will go from A... 1 will...
            97. Wrong..., this is the problem and...
            98. 1 will go from B to C...
            99. For that, um..., this time 3 from B, um..., has to go on C, so...
            100 For that, 2 has to go to A.
            101. For that, 1 has to go back to C, of course.

The generalization of Dsk2s to Dsk3s occurs around lines 98 and 99, and it consists of replacing the constant, "peg A", with a variable.

According to the representation in table 5, rule Dsk3s should be able to handle any situation where the disks blocking the move are stacked on top of the disk to be move. However, there is evidence from later in the protocol that this representation is not quite right. Dsk3s should be able to handle major move 24 and 26, but contrary to prediction, there are signs of an impasse:

[_,1234,5] 140. and then, this time, um..., um, 4 will go to C, so...
            141. 3 goes to A,
            142. 2 goes to B,
            143. and then, 1 will go to A.
                 ....
[123,_,45] 153. So, this time, um..., oh, this time, 3 naturally has to go here, so,
            154. for that, 2 has to go to B.
            153. So 1 will go from A to C,

The impasses at lines 140 and 153 are indicated only by pauses. They might not really be impasses, because many things can cause pauses. If the pauses are not caused by impasses, then there is no problem; the explanation based on the rules in table 5 works fine. On the other hand, if the pauses are caused by impasses, then the explanation is only a little bit off. We could correct it by changing the rules so that they include more contextual information. The mismatch between the contextual information of Dsk3s and the context at lines 153 and 140 causes impasses there, and the impasses trigger further generalization.

So far, we have covered only cases where the blocking disks lie on top of the disk to be moved.

The rules for the other case, where the blocking disks lie on the goal disk's destination, can also be explained by a series of impasses, each one causing a minimal generalization (see table 6, below).

In summary, there seem to be two distinct acquisition mechanisms involved in the learning of the disk subgoaling strategy. The initial version of the strategy seems to have been triggered by reflection and involve some kind of simple inference (deduction? decompiling a chunk?). The subsequent versions are clearly the result of impasse-driven minimal generalization.

4.4. Acquiring the pyramid subgoaling strategy

The acquisition of the pyramid subgoaling strategy has an explanation that is interesting in that the machine learning literature does not yet contain a model that corresponds well with the subject's behavior. At the initial appearance of the pyramid strategy at major move 30, the subject shows no sign of an impasse:

[5,4,123]   178.   Next, 5 has to go to C, so...
                179.   I only need move three blocking disks to...B.
                180.   So, first...I will go from C to B.

Logically, the subject should not suffer an impasse here, because she earlier handled identical situations and should by now have rules sufficient to handling this one without an impasse.

The verbal evidence is too sparse to determine how the pyramid strategy was acquired. However, since the 3-high pyramid is in full view on peg C, I suspect that Anzai and Simon are right in conjecturing that the subject has merely substituted the perceptually more salient concept of "pyramid" for the concept "disk" in the old disk subgoaling strategy. As table 2 shows, a simple substitution of one concept for another is sufficient to effect the change (given that the rules are expressed in English; there are formal representations where simple substitution will also suffice.) Apparently, this substitution of "pyramid" for "disk" does not seem to be triggered by an impasse.

If substitution is what happened at lines 178-180, then the resulting pyramid strategy should be just as general as the disk strategy from which it was formed. Thus, there should be no impasse at its next use, major move 32, because that state, [_,1234,5], was handled by the disk strategy in the preceding episode. As predicted, the verbal evidence shows no sign of an impasse:

[_,1234,5]   188.   It's easy, isn't it?
                189.   5 has already gone to C.
                190.   Next..., 5 was able to move, because...
                191.   A and C were open, right?
                192.   5 is already at C, so...

193.   I will move the remaining four from B to C...
194.   It's just like moving four, isn't it?
195.   So...I will have to move 4 from B to C...
196.   For that, the three that are on top have to go from B to A...
197.   Oh yeah, 3 goes from B to A!
198.   For that, 2 has to go from B to C,
199.   for that, 1 has to go from B to A.
200.   So, 1 will go from B to A.

The subject is not at an impasse. On the contrary, she appears to have two applicable strategies, the disk and pyramid strategies, believes both are correct, and is happily comparing them by running them in synchrony. Line 193 is from the pyramid strategy; line 195 is from the disk strategy; line 196 is from the pyramid strategy; line 197 is from the disk strategy. The remaining lines are from the disk strategy. This segment of protocol is another illustration of the "scientific" approach this subject takes to strategy acquisition. Before adopting the new strategy, she methodically compares it, step by step, with the old strategy.

No single program in the machine learning literature corresponds to the conjectured learning taking place here, because it involves a combination of mutation-based learning and explanation-based learning. Mutation-based learning (DeJong, 1988; Lenat, 1983) makes random changes to the knowledge representation and keeps the mutated representations only if they lead to better overall performance of the system. In this case, a random mutation of the disk subgoaling strategy is tested not by measuring its performance, but by explicitly comparing its operation to that of its predecessor, and thus demonstrating analytically that it is just as correct as its predecessor. The use of analytic as opposed to empirical methods for demonstrating correctness is characteristic of explanation-based learning methods (Mitchell, Keller & Kedar-Cabelli, 1986), so this subject's learning seems aptly described as a combination of mutation-based and explanation-based learning. It would be an interesting implement a machine learning program with both methods in it in order to explore the power of this combination.

5. Discussion

The preceding sections uncovered the lines where strategy acquisition occurred and characterized each occasion in terms of the computational methods that best fit the verbal data. This section discusses what has been learned from this analytic exercise. First, the main result of the analysis is presented, which is an explication of the mapping between behavior and machine learning models. Then, two issues in the skill acquisition literature are reviewed, and the data from this analysis

is brought to bear on them.

5.1. Rule acquisition events

If the preceding analysis is correct, then this subject's strategy acquisition takes place during 11 events, which I call *rule acquisition events*. Table 6 lists them. These events were determined by first finding a set of rules and rule firings (see table 3) that maximizes the fit of the model to the protocol, then searching the protocol near the vicinity of the first use of a rule for signs of unusual cognitive activity, then finding an explanation in terms of machine learning mechanisms that could explain the rule's acquisition as well as the subject's utterances.

```
-----------------------------
Place table 6 about here.
-----------------------------
```

The rule acquisition events were of moderate length, lasting between a few seconds and a couple of minutes, with a mean duration of approximately 75 seconds.[8] If we assume, following Newell (in press), that the cognitive architecture has a 100 millisecond cycle time, then a rule acquisition event takes roughly 750 cycles. Although these timing estimates are too crude to constrain the choice of acquisition mechanism, they do give us a ballpark figure for the amount of computation required to effect a rule change. These particular rule acquisition events are neither the simple application of a single inference rule or schema, nor a gigantic search of a rule space.

In all but one case (the first appearance of the pyramid rule, at lines 178-180), rule acquisition events were accompanied by signs of unusual cognitive activity, such as long pauses, negative comments, or reflective announcements of new insights into the puzzle's structure. The eleven rule acquisition events correspond to major moves 3, 13, 14, 18, 21, 22, 24, 26, 30 and 32. (There are two at major move 14.) As column 4 of table 3 shows, 90% (10/11) of the rule acquisition events are accompanied by extensive non-goal talk. In contrast, only 4% (5/119) of the other moves in the protocol were accompanied by extensive non-goal utterances. The probability of this occurring by chance is less than 0.001 by Chi-squared.[9]

The primary goal of this research was to find out how machine learning methods manifest themselves in human behavior. The answer seems to be that acquisition methods manifest themselves in protocol data in the same way that the classic weak methods of problem solving do. Sometimes the steps in the acquisition method are quite visible. For instance, the Anzai and Simon subject is quite

explicit about comparing the disk and pyramid strategies during the rule acquisition event of lines 192-200. At other times, the execution of a method can only be inferred by a pause followed by the appearance of its results, as in the rule acquisition events where the subject generalized the disk subgoaling strategy.

There is one potential difference between manifestations of acquisition methods and weak methods. When using a problem solving method, subjects tend to announce its results whenever it produces any. Rule acquisition methods produce rules as their results, but one rarely hears as subject announce a rule in recognizable form. About the closest that the Anzai and Simon subject comes to announcing a rule is when she says, "that anyway 2 will have to go to the bottom of C, naturally I thought of 1 going to B." (line 78) I suspect that this subject's behavior is typical, and that new rules, when they are announced at all, will be spoken of in an instantiated form. This explains why it is so difficult to infer rule acquisition events from protocol data. Lacking direct information about the results of acquisition methods, one must use the subject's overt behavior (e.g., overt moves, goal utterances) to infer changes in the problem solving rules, and hence infer the operation of acquisition mechanisms.

The fact that the subject never mentions a rule in abstract form is consistent with Schank's (1982) hypothesis all knowledge is encoded in a highly instantiated form, called cases. This assumption explains why the subject mentions specific disks and pegs. Of course, the data equally well support the hypothesis that subject finds it easier to use specific language to describe rules that are encoded mentally as general abstractions. It takes fewer words to say "peg C" than "the peg I intend to move the disk to." There is, of course, a compromise position wherein the subject has both general and specific encodings, as suggested by studies of concept formation (Elio & Anderson, 1981) and by the gradual generalization of the disk subgoaling strategy.

The conclusion that strategy acquisition takes place in a series of rule acquisition events is not shared by Schoenfeld, Smith and Arcavi (in press), who analyzed a protocol from a student who was learning about graphing linear equations. They report (in draft of May 1989, pg. 10-11):

> A main purpose of our study was to provide a detailed exegesis of one student's learning -- a microgenetic analysis of her cognitive change, describing how her understanding of the domain evolved. In many ways, it seems natural to look for "learning events" as sites of cognitive change. Suppose, for example, that at time T1 the student enters into an interaction with either the computer or the tutor. The student begins with a particular knowledge state, say KT1. She performs some action, gets some feedback, interprets that feedback, and arrives at a new knowledge state, KT2. This sequence (typically taking place over a time frame of seconds, at most a few minutes) results in a relatively small cognitive change. Over time, of course, these changes build up: The sequence of micro-changes should result in the kind of "macro" learning that characterizes the large "beginning to end" changes we saw in [the

subject]....

> Though we combed the tapes carefully looking for learning events of the type just described, we report having found remarkably few. It appears that for semantically rich domains simple "learning is adding knowledge to the knowledge base" models or straightforward "adding productions to the production system" models of learning (see, e.g., Klahr, 1978) do not do justice to the complex, unstable, and non-monotonic aspects of human learning. [The subject's] learning was slow, organic, and often retrogressive (i.e., old knowledge died hard, often recurring even after it appeared to have been "replaced").

The authors go on to suggest that learning could be more faithfully captured by some kind of connectionist network, presumably because connectionist networks produce slow learning with plenty of retrogressions.

In most respects, the present analysis agrees with that of Schoenfeld et al. In both cases, rule acquisition events are infrequent. Eleven were found in this 90 minute protocol, so Schoenfeld et al. should have found only 22 in their 180 minute protocol. They appear to have found less than that, but this is probably due to the differences in the analytic techniques employed. Schoenfeld et al. did not build a computer simulation of the protocol. Without a simulation, it is difficult to detect when the subject is using different rules than she did before. One is forced to look for the signs of unusual verbal behavior that sometimes, but not always, accompany rule acquisition events. Still, if their subject was as vocal as the Anzai and Simon subject, then 90% of their subject's rule acquisition events should have been accompanied by unusual verbal behavior. If 22 events actually occurred, then Schoenfeld et al. should have found 20 rule acquisition events, which is more than they appear to have found. However, the nature of their training situation probably hid the verbal behavior that often accompanies rule acquisition events. Their subject was learning from a human tutor, and most of her learning seems to have occurred while they were conversing. Because the subject was not really giving a concurrent protocol, the transcript may often lack verbal signs of the rule acquisition events that occurred. If only 10% of their subject's rule acquisition events were accompanied by unusual verbal behavior, then they should have found 2 rule acquisition events, which is probably close to what they actually found.

Both the Schoenfeld et al. protocol and this protocol showed quite a bit of retrogression. As table 3 shows, the initial firing of a rule did not herald a consistent use of it on every subsequent occasion. Instead, there was always at least one and sometimes many subsequent moves where old rules were fired in place of new one. The gradual or retrogressive nature of strategy acquisition has been noted before (Kamiloff-Smith & Inhelder, 1974; Lawler, 1981; Kuhn & Phelps, 1982; Siegler & Jenkins, 1989), and the next section gives a thorough discussion of the phenomenon. In the Tower of Hanoi protocol, it

appears to be caused mostly by failure to retrieve the new rules from memory.

The main difference between the Schoenfeld et al. study and this one is that different conclusions are drawn from similar data. Because I was able to define rule acquisition events for 100% of the subject's learning and align 90% of them with periods of unusual activity in the protocol, I conclude that rule acquisition events exist and that we can infer much about acquisition mechanism from studying them, including, for example, why learners often regress to older strategies. Schoenfeld et al. found "remarkably few" acquisition events and conclude that some invisible, perhaps connectionist acquisition mechanism is responsible for the subject's learning. It seem more plausible to me that their subject did experience a variety of acquisition events, but their analytical methods were not appropriate for locating them.

Although we have attained the main objective of the analysis, finding out how acquisition methods mapping onto human behavior, there are several important issues raise by the analysis. They next few subsections discuss them.

5.2. Why do people forget what they discover?

It is often thought that discoveries not only occur suddenly, with a flash of insight or an inductive leap, but that they produce an indelible memory in the mind of the discoverer. This belief is part of the enduring belief among educators that discovery learning is an effective instructional technique. This section shows that there may, in fact, be nothing special about the durability of memory traces created by rule discovery processes.

The Anzai and Simon protocol is an example of discovery, because the subject acquired new strategies without help from the experimenter. However, the subject's learning did not have the indelibility that supposedly characterizes discovery learning. With all three rules, the initial firing of the rule was followed by a missed opportunity (see table 3). Later, the rules (especially the pyramid strategy) came to be fired whenever conditions were appropriate, but their initial firing pattern was intermittent.

This same pattern of intermittent initial usage has been found by other investigators (Kamiloff-Smith & Inhelder, 1974; Lawler, 1981; Kuhn & Phelps, 1982; Schoenfeld, Smith & Arcavi, ress; Siegler & Jenkins, 1989). In Kuhn and Phelp's (1982) study, for instance, 7 subjects discovered a correct

solution strategy, but only one used this strategy consistently after the initial discovery of it. The authors state that this subject "was the striking exception. The more characteristic pattern was a much more gradual acquisition, with a sustained period during which more advanced strategies were used in conjunction with less advanced ones." (Pg. 26)

Since the Tower of Hanoi is such a simple task domain, we are in a good position to uncover the mechanisms responsible for this pattern of gradually increasing usage. There appear to be two.

One is competition between old and new strategies (Siegler, 1986). In the case of the subgoaling strategies, the old strategy is rule 2blk, which moves a 2-high pyramid out of the way when it is blocking a desired move. At the beginning of episode 3, the subject appears to adopt a policy of not using this rule in apparent attempt to gain further understanding into the puzzle and a less ad hoc strategy for solving it (see the section of scientific strategy acquisition). However, at several points, the subject uses rule 2blk instead of rules Dsk or Pyr (see table 3). Twice (major moves 18 and 21), she catches her slip and redoes the move's reasoning with the subgoaling rules. These retrieval failures probably have two causes. The new rules were less familiar because they had been used less often. Visual cues may also contribute to the unintentional retrieval of 2blk over the subgoaling rules. Rule 2blk is concerned with moving the 2-high pyramid out of the way. On 6 of its 7 applications, the 2-high pyramid was visually salient because it was sitting either alone on a peg or on top of the largest disk. In short, it appears that ordinary mnemonic variables, such as familiarity and visual cuing, explain part of the the pattern of gradually increasing frequency observed in this protocol and, most likely, other studies as well.[10]

A second mechanism for explaining gradually increasing frequency shows up in the acquisition of the disk subgoaling strategy. The pattern of impasses and non-impasses is accurately explained by assuming the strategy was first learned in a highly specific form then generalized only when necessary. Rule generalization was originally thought to be part of the cognitive architecture (Anderson, 1983), but recent experiments by Anderson and his colleagues (Anderson, 1987) suggests that the subject has conscious control over the process. In this protocol, conscious control would be consistent with the subject's methodical investigation of the solution space. When inducing a hypothesis, it is a good heuristic and a common scientific practice to avoid overinterpreting the data. This would predict a pattern of conservative generalization of a (deliberately?) overspecific initial version of the hypothesis, which is exactly what is observed.

Conservative generalization processes, whether conscious or not, can explain some patterns of increasingly frequent usage, although one must adopt an extra assumption to do so. The assumption is that a person is only willing to do a small amount of generalization at a time. For instance, the subject might be willing to changes two constants to variables, but be unwilling to change ten. If the current version of the rule does not match the current situation, and a small amount of generalization will not allow it to do so, then some other rule is chosen instead of this rule. This could cause patterns of gradually increasing rule usage as the rule gradually becoming more general. Some of the earlier missed opportunities in the development of the disk subgoaling strategy (e.g., major move 17) might be due to such a process. The failure of rule 4B to apply at major moves 4, 20 and 28 might also be due to a lack of generality.

However, whenever a generalization-based explanation fits the retrogression facts, so does an explanation based on retrieval failure. Hence, it is not possible to argue conclusively from these data that people will refuse to generalize too fast. On the other hand, retrieval failure does have good support in the data. There are major moves (27, 29 and 31) where only retrieval failure will explain the missed opportunities because the new rules in question are already completely general. So we have firm evidence that retrieval failure is a factor in the gradual increase in rule firing, but only confounded evidence for the participation of conservative generalization.

5.3. Does all learning occur at Impasses?

An old and controversial hypothesis is that failures are particularly important stimulus for learning (Schank, 1982). One version of this hypothesis is that a particular kind of failure, called an *impasse*, is the source of almost all procedural learning (VanLehn, 1988; Laird, Rosenbloom, & Newell, 1986). Although the term "impasse" can only be properly defined relative to a given cognitive architecture, its approximate meaning is that the architecture cannot decide what to do next given the knowledge and the situation that are its current focus of attention.

Architectures that support impasses have an automatic mechanism that causes a shift of attention -- the impasse itself becomes the problem to be solved and that causes new knowledge to become relevant, knowledge about how to solve impasses. If this knowledge suffices for resolving the impasse, the architecture stores in long-term memory a summary of the impasse and its resolution. The next time the kind of situation that caused the old impasse occurs, the architecture can recall how it resolved it the

last time and use this memory to guide its actions. It has learned what to do at these kind of situations, so it no longer reaches an impasse there.

Because impasses are defined relative to a given architecture, there is substantial differences in what counts as evidence for an impasse in human behavior. The Soar architecture, like ACT* (Anderson, 1983), is primarily a model for memory and attention (Newell, 19??). It's basic cycle time is intended to be about the same order of magnitude as simple memory access for humans -- around 100 milliseconds. The only way to augment Soar's long term memory is via an impasse. People are known to update long term memory nearly continuously, so a proper Soar model should have impasses at least every few seconds. In contrast, Sierra and its successors (VanLehn, 1987; VanLehn, 1990; VanLehn & Ball, in press) are intended as a model of planning and procedure following. Their basic cycle time is intended to be the same order of magnitude as a person's decision about what to do next while problem solving -- typically 1 to 10 seconds. Their impasses are intended to correspond to occasions where a subject really is stuck while problem solving, which often show up in protocols as long pauses or complaints (e.g., "yes, yes, now it gets difficult. Yes, it's not that easy...."). These two kinds of impasses are sometimes distinguished coloquially by the terms "big impasses" (Sierra) and "little impasses" (Soar).

In earlier work (VanLehn, 1988), I claimed that big impasses are the triggers for all rule acquisition events. Of course, big impasses do not occur frequently enough to explain all updates to long term memory, so I also assumed that other acquisition mechanisms were responsible for other kinds of learning (e.g., noticing things on your commute into work, the power law of practice, etc.). On the other hand, the Soar group has claimed that little impasses are the sole triggers for all kinds of learning (Newell, 19??). This is clearly a much different claim than mine and not as easily tested with protocol data.

In order to test my impasse-driven learning hypothesis, we can use the verbal criterion for impasses mentioned earlier (negative comments, long pauses). Applying this criterion to the protocol yields a collection of 11 impasses (lines 12, 17, 20, 32, 53, 80-81, 97, 119, 128, 140, and 153). The impasse at line 53 seems to have been caused by the experimenter's request for an explanation of a move, so it will be ignored. The impasses at lines 17 and 20 occurred at the end of episode one, where the subject already knew she was in trouble. The remaining 8 impasses occurred at major moves. Since the protocol lasted 90 minutes, this works out to about one impasse every ten minutes.[11] Thus,

these verbal signs are indicative of big, Sierra-sized impasses. They have little to do with small, Soar-sized impasses.

Given this rough definition of impasses, only 73% (8/11) of the rule acquisition events in this protocol are impasse-driven (see 3). There are no verbal signs of impasses at the other three rule acquisition events. When evaluated in this manner, my impasse-driven learning hypothesis is false, even though headed in the right direction.

One could attempt to salvage the impasse-driven learning hypothesis by arguing that impasses occurred at all the rule learning events, but on three of them the subject did not exhibit any verbal signs of an impasse. I do not think that the notorious incompleteness of verbal protocols should be invoked in this case. At the three rule acquisition events in question (lines 78, 178-180, and 193-200), there is no reason for the subject to be stuck. Her lack of impasses at similar, earlier situations argues that she had the knowledge needed to continue without pause at those puzzle states. Furthermore, there are plausible non-impasse explanations for the triggers of these rule acquisition events, as discussed earlier.

The lack of impasses at some rule acquisition events is consistent with evidence from Siegler and Jenkin's study (1989). Siegler and Jenkins conducted a longitudinal study of the development of the addition strategies of 4 and 5 year old children. The findings of interest here concern the discovery of the Min strategy. Suppose the problem is to add 3 plus 4. In one version of the Min strategy, the child raises fingers for the smaller addend (3), then touches each of them while counting on from the larger addend (4). The Min strategy is much more efficient than its predecessor, the Sum strategy. The Sum strategy counts out sets for each addend, then counts up the union of those two sets. In solving 3+4, the child might count out three fingers on one hand, then four fingers on the other, and finish by counting all the fingers up.

When children began the experiment, they often used the Sum strategy along with retrieval from memory and other strategies. However, they did not use the Min strategy. Using protocol and timing data, Siegler and Jenkins established when the first use of the Min strategy occurred. They report,

> Two consistent patterns characterized performance just before the discoveries: long solution times and appearance of the shortcut-sum strategy. On the trial immediately before discovery of the Min strategy, average solution times were twice as long as the average for the experiment as a whole. The long solution times might be interpreted as indicating that the problems were very difficult. This was not the case, however; they were quite representative of the total set of problems that children encounter in the study. Further, the same child often had solved the same problem much more rapidly and without any

obvious difficulty earlier in the experiment. [pg. 101]

It appears that 6 of the 8 students learned the Min strategy via reflection rather than via an impasse. Of the remaining two subjects, one did not discover the Min strategy, and the other discovered it while trying to solve the problem 1+24, a problem designed to cause an impasse for the Sum strategy.

Siegler and Jenkins also showed that the initial use of the Min strategy is not followed by an exclusive reliance on it. Rather, it seems almost to be forgotten about, for it is used only sporadically up to the session where subjects were first given "challenge" problems, such as 2+23, that should cause the Sum strategy to reach an impasse. Of the 5 subjects who had already used the Min strategy at least once before receiving the challenge problems, 4 began to use it more frequently. Of those 4, 3 increased their use of the Min strategy on ordinary problems as well as on challenge problems. Those who had not yet discovered the Min strategy continued to use the Sum strategy and did not discover the Min strategy until much later, if at all.

Siegler and Jenkin's finding~ are strikingly parallel to those from the Tower of Hanoi. In the Tower of Hanoi study, only 1 of the 3 initial rule acquisition events was triggered by an impasse (see table 6). In the Siegler and Jenkins study, only 1 of the 7 subjects discovered the Min strategy in a context where an impasse was likely. However, in both studies, subsequent generalization of new strategies does seem to be driven by impasses.

There is a plausible reason why the initial acquisition events were not impasse-driven in these two studies. In both these studies, the subjects had perfectly good strategies for solving problems. They had the Sum strategy for addition and the Selective Search strategy for the Tower of Hanoi. During ordinary usage, these strategies could not reach an impasse. In order to reach an impasse, the subject would have to do something special, such as deliberately ignoring the existing strategy. But why would the subject do that? In the Tower of Hanoi, the discovery of both subgoaling strategies seems to have come, ultimately, from some kind of "noticing" activity (cf. Ruiz & Newell, 1989). The subject "noticed" that there was mathematical structure in her episode 2 solution to the puzzle, so she set up the experiment that eventually led to the discovery of the disk subgoaling strategy. The subject "noticed" the coincidence of the perceptually salient pyramid concept with the disk concept used in her subgoaling strategy, so she substituted one for the other. Granted, the evidence for this "noticing" activity is slim, but that is to be expected given the notorious difficulty that psychology has had in isolating the processes involved in insight.

The picture painted by these speculations is that there are basically three kinds of triggering conditions for rule acquisition events. The most common is the kind of impasse where subject cannot decide what to do next. The second class consists of impasses where the subject is deliberately ignoring knowledge that would otherwise allow her to decide what to do. The third class of triggering conditions is a miscellaneous collection that includes perceptual noticing (pyramid case), episodic noticing (disk case), deliberate reflection on past solutions and perhaps other mechanisms as well. It appears that the ecology of rule acquisition events is much more complicated than anyone suspected before.

6. Conclusions

6.1. Summary

The main objective was to find out how strategy acquisition processes would manifest themselves in human behavior. A model was fit to the protocol of a subject who solved the Tower of Hanoi puzzle several times. Rules were invented and their exact firing sequences were adjusted in order to maximize the number of overt moves and goal utterances covered by the model. On this basis, we could infer when the subject first started to use new rules. A *rule acquisition event* was defined to be a time when the model constructs or modifies a rule. The model for this protocol produced 11 rule acquisition events. Because the model was fit so closely to the protocol, it was possible to locate these events in the protocol. This allowed the observation of the processes of strategy acquisition in action.

It was found that on 90% of the rule acquisition events in this protocol, the subject gave some signs of unusual cognitive activity, such as a long pause, or a negative comment (e.g., "Yes, Yes, now it gets difficult...") or a reflective comment (e.g., "That anyway, 2 will have to the bottom of C, so naturally I thought of 1 as going to B."). Apparently, the subject interrupts her ordinary pursuit of the problem in order to think briefly about her solution method. While doing so, the protocol shows the normal signs of a goal directed process, with the usual incomplete statements of goals, results and difficulties. Curiously, this subject never mentioned rules in a general or abstract form even though she shows visible excitement at some of the acquisition events, as if she knew that she had discovered something general about the puzzle. In short, it appears that 90% of the acquisition in this protocol manifested itself as serial, goal-directed processing that interrupts the flow of ordinary problem solving. This is the main finding.

It was also found that the subject did not consistently use a rule after having discovered it. This contradicts the common sense model of discovery learning, which holds that the products of discovery are written indelibly in the mind by a blinding flash of insight. This subject seemed to forget temporarily the rules she had discovered, so her use of the new rules increased gradually. Similar findings have been reported by others (Lawler, 1981; Kuhn & Phelps, 1982; Siegler & Jenkins, 1989; Schoenfeld, Smith & Arcavi, ress). Because the Tower of Hanoi is such a simple task domain, it was possible to develop a fine-grained analysis of the subject's failures to apply the new rules. This analysis showed that this subject's neglect was caused by retrieval failure and perhaps also by possession of overly specific rules.

Some years ago, I found that correct subtraction procedures and most buggy subtraction procedures could be explained by a trio of acquisition methods, all of which were driven by impasses (VanLehn, 1988). In a fit of enthusiasm, I claimed that *all* procedural skill acquisition occurred at impasses. This present study showed that this claim was a bit too strong. Only 73% of the rule acquisition events are driven by impasses. Of course, that is still a lot, so the impasse-driven learning hypothesis is still a reasonable summary of human behavior.

It was also found that impasse-driven learning was less common among *initial* rule acquisition events (33%) than among subsequent rule acquisition events (87%). Siegler and Jenkins (1989) report a similar finding. The explanation, in both cases, seems to depend on the fact that in these protocols the subjects were improving strategies that already produced optimal, error-free solutions to problems. With such good strategies, the subjects would not normally encounter impasses, so it is to be expected that impasses would not drive their initial rule discoveries. In other learning situations, subjects might improve their strategies because their old ones do not always give correct answers. In such cases, a greater percentage of the rule acquisition events will be caused by impasses. As arithmetic is probably a case where errors motivate the learning, it makes sense that most of its rule acquisition events would be impasse-driven.

6.2. Speculations

Protocol analysis tends to give the analyst strong intuitions about the cognitive architecture, even though the data themselves are so remote from architectural mechanisms that they provide tangential support at best. This protocol suggests the following hypotheses:

- Rule acquisition methods are just like problem solving methods, except that they focus on the problem of improving a problem solving method.

- Rule acquisition methods require the subject's attention in order to function. Thus, they temporarily supplant ordinary problem solving, which also requires the subject's attention. It is not the case that rule acquisition methods run in parallel with ordinary problem solving, say, at a subsymbolic level.

- Whereas the typical problem solving method deposits in memory some decision about what actions to take, the typical rule acquisition method deposits a new rule in memory. Whether or not the subject will remember this rule on subsequent occasion depends on the usual host of mnemonic factors. For instance, the rule may have to be (re-)constructed many times before it s familiar enough to be used retrieved reliably.

Although these claims go well beyond the data found in this single protocol, it is worth putting them on the table, since they do provide a plausible interpretation of the evidence.

In an early report on this work(VanLehn, 1989b), I followed Schoenfeld et al. (in press) in using the term "learning event" to refer to changes in the subject's procedural knowledge. Given the present analysis, "learning" is an inaccurate label for those events. Psychologists use the word "learning" for the storage of any sort of information in memory including, for instance, specific puzzle states and moves. The 11 events studied here involve the onset of a durable new behavior, which, according to the speculations above, is caused by inventing a new rule and storing it in memory. It is important to emphasize that inferring a rule (i.e., rule acquisition) is different from learning a rule (i.e., storing it in a retrievable form). Thus, for instance, Soar's chunking mechanism (Laird, Newell, & Rosenbloom, 1987) is probably a good model of learning (memory storage) but it is not a model of rule acquisition. In Soar terms, rule acquisition would be a special problem space (or spaces) whose operators explicitly construct rules. The distinction between rule acquisition and rule learning is even clearer in ACT* (Anderson, 1983). A rule acquisition method corresponds to productions that construct a semantic net representation of a new rule. This network is stored in working memory, and may, with a certain probability, become permanent. Thus, the act of inferring a new rule is distinct from the act of learning it. In short, the term "learning event" is a misnomer, and "rule acquisition event" is much better.

The success of the overall approach taken by the subject suggests a new directions for machine

learning research. This particular subject pursued a scientific investigation of strategies for solving the puzzle. She appears to have explicitly rejected old strategies in order to force the development of new ones. She designs and carries out a programme of experimentation, during which she first conjectures a new strategy then methodically tests and generalizes it. No existing machine learning program exhibits such a scientific approach to strategy acquisition. The current favorite approach to strategy acquisition has the machine take the role of an apprentice (Mitchell, Mahadevan & Steinberg, 1985). These learning apprentices wait passively for the user to give it problems and examples that stretch its understanding. The success of the subject in this experiment suggests that a better metaphor for machine learning might be the bright young scientist who is always coming into your office and pestering you with questions and requests for research projects.

The fact that this subject adopted a scientific approach to strategy acquisition is not the only thing that makes her unusual. Although many subjects can master the Tower of Hanoi in about the same length of time as this subject, this subject appears to be one of the best in the literature (Ruiz & Newell, 1989). The fact that a good student adopts a such rational approach to acquiring knowledge fits nicely with a finding from Chi, Bassock, Lewis, Reimann and Glaser (1989). These authors discovered that students who explained physics examples to themselves were better learners than students who merely read the examples and paraphrased them. Self-explanation involves working the problem stated in the example, and checking one's solution against the solution given in the book. This is a rational way to check that one understands the subject matter, since it allows one to localize knowledge deficits. If one cannot connect two adjacent lines in the solution in the same way that the book does, then one has a smaller space to search for the knowledge deficit than would have to search if one knew only that one's solution did not reach the same final answer as the book's. Hence, self-explanation is an excellent way to utilize the given instruction. Pirolli and Bielaczyc (1989), working with Lisp students, found the same correlation between self-explanation and learning.

Self-explanation and the kind of scientific strategy acquisition observed in the Tower of Hanoi subject are both based on the same premise, that one's knowledge base is potentially inadequate and may need improvement. The difference between the two acquisition methods is determined solely by the instructional setting. Self-explanation is a good method (if not optimal) when the instruction includes solved example problems. Scientific strategy acquisition is a good or optimal method if there is no instruction at all. In short, it may be that good students are good because they assume that they are

ignorant, they want to become less ignorant, and they know how to solve the ignorance problem in a variety of instructional settings.

Acknowledgments

Appendix

This appendix presents the protocol analyzed in this paper, which is take from Anzai and Simon (1979). Most of the protocol appears in table 7, which lists the subject's comments during her 3 complete solutions of the five disk puzzles. The rows of the table correspond to movements of individual disks, the columns correspond to episodes, and the cells contain the subject's comments as she planned and made the move. There are two segments of protocol that are left out of this table: episode 1 and the solution of the smaller puzzles that occurred as part of episode 3. For completeness, these are listed below.

1   I'm not sure, but first I'll take 1 from A and place it on B.

2   And I'll take 2 from A and place it on C.

3   And then, I take 1 from B and place it on C.

       (Experimenter: If you can, tell me why you place it there?)

4   Because there was no place else to go, I had to place 1 from B to C.

5   Then, next, I placed 3 from A to B.

6   Well..., first I had to place 1 from B, because I had to move all disks to C.

       I wasn't too sure though.

7   I thought that it would be a problem if I place 1 on C rather than B.

8   Now I want to place 2 on top of 3, so I'll place 1 on A.

9   Then I'll take 2 from C, and place it on B.

10   And I'll take 1 and... place it from A to B.

11   So then, 4 will go from A to C.

12   And then..., um..., oh..., um...

13   I should have placed 5 on C. But that will take time. I'll take 1...

(Experimenter: If you want to, you can start over again. If you are going to do

that, tell me why.)

14   But I'll stay with this a little more...

15   I'll take 1 from B and place it on A.

16   Then I'll take 2 from B to C.

17   Oh, this won't do.

18   I'll take 2 and place it from C to B again.

19   And then, I'll take 1, and from A...

20   Oh no! If I do it this way, it won't work.

21   I'll return it.

22   Ok?

23   I'll start over. (Experimenter: Go ahead.)

24   If I go on like this, I won't be able to do it, so I'll start over again.

<lines 25 to 70 are in table 5>

71   I wonder if I've found something new.

72   I don't know for sure, and little ones will have to go on top of big ones...

big ones can't go on top of little ones, so first, bit by bit,

C will be used often before 5 gets there.

73   And then, if 5 went to C, next I have to think of it as 4 to go to C...

74   This is my way of doing it...

75   Can I move it like this?

76   First, if I think of it as only one disk, 1 could go from A to C, right?

77   But if you think of it as two disks, this will certainly go as 1 from A

to B and 2 from A to C, then 1 from B to C.

78   That...that anyway 2 will have to go to the bottom of C, naturally I

thought of 1 going to B.

79   So, if there were three..., yes, yes, now it gets difficult.

80   Yes, it is not that easy...

81   ...this time, 1 will...

82   Oh, yes, 3 will have to go to C first.

83   For that, 2 will have to go to B.

84   For that, um..., 1 will go to C.

85   So, 1 will go from A to C, 2 will go from A to B, 1 will go back from C to

       B, I'll move 3... That's the way it is!

86   So, if there were four disks, this time, 3 will have to go to B, right?

87   For that, 1 will have to stay at C, and then,

       for that, 1 will be at B.

88   So 1 will go to B.

89   And then, 2 will go from A to C.

90   And then, 1 will go back to C from B.

91   And then, 3 will move from A to B.

92   And then, I will move 1 from C to A.

93   And then, first, I will move 2 from C to B.

94   And then, I will move 1 from A to B,

95   and then, 4 from A to C.

96   And then, again this will go from A... 1 will...

97   Wrong..., this is the problem and...

98   1 will go from B to C...

99   For that, um..., this time 3 from B, um..., has to go on   so....

100   For that, 2 has to go to A.

101   For that, 1 has to go back to C, of course.

102   And then, 2 will go from B to A,

103   and then, 1 will go from C to A,

104   and then, 3 will go from B to C.

105   So then, 1 will go from A to B.

106   2 will go from A to C,

107   and then, 1 will go from B to C. (Experimenter: All right.)

<The rest of the protocol is in table 5>

#	Subgoaling	Episode 2	Episode 3	Episode 4
1	5C 4B 3C 2B 1C	...1C	5C 4B 3C 2B 1C	5C 4B 3C 2B 1C
2	" " " "	2B	2B	2B
3	" " " 1B	1B	1B	1B
4	" " "	3C	3C	3C
5	" " 2C 1A	...4B..1A	4B 1A 4B 2C 1A	4B 1A
6	" " "	2C	2C	2C
7	" " 1C	1C	1C	1C
8	" "	4B	4B	4B
9	" 3B 2A 1B	..1B	5C 3B 2A 1B	5C 3B 2A 1B
10	" " "	2A	2A	2A
11	" " 1A	1A	1A	1A
12	" "	3B	3B	3B
13	" 2B 1C	1C	1C	1C
14	" "	2B	2B	2B
15	" 1B	1B	1B	1B
16	"	5C	5C	5C
17	4C 3A 2C 1A	....1A	4C 3A 2B 1A	4C 3A 2C 1A
18	" " "	2C	2C	2C
19	" " 1C	1C	1C	1C
20	" "	3A	3A	3A
21	" 2A 1B	1B	1B	2A 1B
22	" "	4C 2A	2A	2A
23	" 1A	1A	1A	1A
24	"	4C	4C	4C
25	3C 2B 1C	3C 1C	3C 2B 1C	3C 2B 1C
26	" "	2B	2B	2B
27	" 1B	1B	1B	1B
28	"	3C	3C	3C
29	2C 1A	1A	1A	2C 1A
30	"	2C	2C	2C
31	1C	1C	1C	1C

Table 1: The subject's goals compared to the recursive subgoaling strategy

---

Table 2: Abbreviations and descriptions of rules

Rules present throughout the protocol

- *Top-level* The top level goals are to first get disk 5 to peg C, then get disk 4 to peg C, then disk 3, disk 2 and finally disk 1.

- *Not-twice* It is illegal to move the same disk on consecutive moves.

- *Forced-move* If there is only one legal action, then do it.

- *Put-1-on-2* If there are multiple legal actions, but one of them is to put disk 1 on top of disk 2, thus forming a 2-high pyramid, then do that action.

Initially present rules that disappeared

- *SSS* The Anzai and Simon selective search strategy, plus some special strategies that are applied on the initial moves of episodes 1 and 2 (see Anzai and Simon, 1979, for a description of the special strategies).

- *1blk* If the goal is to move a disk from one peg to another, and there is a single disk blocking the move, then get the blocking disk to the peg that is not involved in the move.

- *2blk* If the goal is to move a disk from one peg to another, and the 2-high pyramid (i.e., disks 1 and 2) is on one of those pegs, then move disk 1 to the other peg (thus freeing disk 2 to move to the peg not involved in the move).

Rules acquired during the protocol

- *4B* Before attempting any of the top level goals, try to get disk 4 to peg B.

- *Dsk* (The Anzai and Simon disk subgoaling strategy.) If the goal is to get a disk from one peg to another, and there are some disks blocking the move, then get the largest blocking disk to the peg that is not involved in the move.

- *Pyr* (The Anzai and Simon pyramid subgoaling strategy.) If the goal is to move a pyramid from a peg to another peg, then get the next smallest pyramid to the peg that is not involved in the move.

Table 2: An analysis of the major moves

Maj. Move	Protocol			Initial rules			Learned rules		
	State	Goal utterances	Extra	SSS	1blk	2blk	4B	Dsk	Pyr.
1	12345, _, _	1B		1			0	0	0
2	45, 3, 12	2B, 1A		1	1		0	0	0
3	5, 123, 4	5C, 1A	14·	1			0	0	0
4	12345, _, _	1C	33+	1			0	0	0
5	45, 12, 3	4B, 1A	50·	0		1	1	0	0
6	5, 4, 123	1B	14+	1				0	0
7	125,34,_	1C		0		1		0	0
8	_, 1234, 5	1A	39·	1				0	0
9	3, 4, 125	1B		0		1		0	0
10	123, _, 45	3C, 1C		0		1		0	0
11	_,12,345	1A		0	1			0	0
12	1, _, _	1C	98+						
13	12, _, _	1B	22+	0	1				
14	123, _, _	3C, 2B, 1C	30·	0		0		1	0
15	_, 12, 3	quits		0	0			0	0
16	1234, _, _	3B, 2C, 1B		0				1	0
17	4, 3, 12	1A		0		1		0	0
18	_, 123, 4	1C		0		1		0	0
"	"	3B, 2A, 1C	30·	0		0		1	0
19	12, _, 34	1B		0	1			0	0
20	12345, _, _	5C, 4B, 3C, 2B, 1C		0			0	1	0
21	45, 12, 3	4B, 1A	22·	0		1	1	0	0
"	"	4B, 2C, 1A		0		0	1	1	0
22	5, 4, 123	5C, 3B, 2A, 1B	17·	0				1	0
23	125,34,_	1C		0		1		0	0
24	_, 1234, 5	4C, 3A, 2B, 1A	7·	0				1	0
25	3, 4, 125	1B		0		1		0	0
26	123, _, 45	3C, 2B, 1C	8·	0		0		1	0
27	_,12,345	1A =		0	1		~	0	0
28	12345, _, _	5C, 4B, 3C, 2B, 1C		0			0	1	0
29	45, 12, 3	4B, 1A		0		1	1	0	0
30	5, 4, 123	5C, 3pB, 1B		0				0	1
31	125,34,_	1C		0		1		0	0
32	_, 1234, 5	4pC,4C,3pA,3A,2C,1A	39+	0				1	1
33	3, 4, 125	2pA, 1B		0		0		0	1
34	123, _, 45	3pC, 2pB, 1C		0		0		0	1
35	_,12,345	2pC, 1A		0	0			0	1

Table 4: Protocol at the initial firing of the three acquired rules

The initial firing of rule 4B

[45,12,3]	30.	And so I'll place 1 from B... to C.
	31.	Oh Yeah! I have to place it on C.
	32.	Disk 2... no, not 2, but I placed 1 from B to C... Right?
	33.	Oh, I'll place 1 from B to A. (Experimenter: Go ahead.)
[145,2,3]	34.	Because... I want 4 on B, and if I had placed 1 on C from B, it wouldn't have been able to move.

The initial firing of rule Disk

[123,_,_]	79.	So, if there were three..., yes, yes, now it gets difficult.
	80.	Yes, it's not that easy...
	81.	...this time, 1 will...
	82.	Oh, yes, 3 will have to go to C first.
	83.	For that, 2 will have to go to B.
	84.	For that, um..., 1 will go to C.

The initial firing of rule Pyr

[5,4,123]	178.	Next, 5 has to go to C, so...
	179.	I only need move three blocking disks to...B.
	180.	So, first...I will go from C to B.

Table 5: Increasingly general versions of the disk subgoaling strategy.
Underlines indicate the generalizations.

name   description

Dsk0   If the goal is to move disk 2 from peg A to peg C,
and disk 1 lies on top of disk 2,
then try to move disk 1 to the peg that is not involved in the desired move.

Dsk1s   If the goal is to move disk X from peg A to peg C,
and disk Y lies on top of disk X,
then try to move disk Y to the peg that is not involved in the desired move.

Dsk2s   If the goal is to move disk X from peg A to some target peg T,
and disk Y lies on top of disk X,
then try to move disk Y to the peg that is not involved in the desired move.

Dsk3s   If the goal is to move disk X from some source peg S to some target peg T,
and disk Y lies on top of disk X,
then try to move disk Y to the peg that is not involved in the desired move.

Dsk1t   If the goal is to move disk 4 from peg A to peg B,
and disk 2 is on peg B on top of the place that disk 4 would go,
then try to move disk 2 to the peg that is not involved in the desired move.

Dsk2t   If the goal is to move disk X from peg A to some target peg T,
and some disk Y is on peg T on top of the place that disk D would go,
then try to move disk Y to the peg that is not involved in the desired move.

Table 6: Rule acquisition events in the Tower of Hanoi protocol	
Lines	Description
11-13	At the state [45,123,_], the subject reaches an impasse when she realizes that she is forced to put disk 4 on peg C, but that will block the top level goal of moving 5 to C. She infers rule 4B, that disk 4 must be on peg B before moving 5 to C.
78	Just after solving the 2-high puzzle, the subject reflects, saying "anyway, 2 will have to go to the bottom of C, naturally I thought of 1 going to B." She infers rule Dsk1s, a very specific version of the disk subgoaling strategy.
79-83	At the state [123,_,_], the subject reaches an impasse because she is trying to solve the puzzle without using her older rules (SSS and 2blk). She resolves the impasse by substituting variables for the constants that refer to disks 1 and 2 in Dsk1s, thus producing Dsk2s.
84	With the goal of getting disk 2 to peg B in state [123,_,_], the subject reaches an impasse and substitutes a variable, T, for a constant, peg C, in rule Dsk2s, thus producing rule Dsk3s. Dsk3s is "If the goal is to move a disk X from peg A to some target peg T, and disk Y lies on top of disk X, and disk Y is the next size smaller than disk X, then try to move disk Y to the peg that is not involved in the desired move."
96-100	At the state [_,123,4], the subject reaches an impasse. The impasse is resolved by substituting a variable, S, for a constant, disk A, in rule Dsk3s.
119-122	At the state [45,12,3], the subject reaches an impasse. After applying rule 2blk to find out the right move, the subject forms rule Dsk1t.
128-130	At the state [5,4,123], the subject reaches an impasse, generalizes rule Dsk1t by changing constants to variables, and thus produces rule Dsk2t.
140-141	At the state [_,1234,5], the subject pauses briefly before applying rule Dsk3s. This may be an instance of an impasse followed by generalization.
153-154	At the state [123,_,45], the subject pauses briefly before applying rule Dsk3s. This might also be an instance of impasse-driven generalization.
178-180	At the state [5,4,123], the subject refers to "three blocking disks" with no signs of an impasse. Some mutative process, perhaps perceptually based, has modified rule Dsk2t by substituting the concept of pyramids in place of the concept of disks.
188-200	At the state [_,1234,5], the subject actively compares her disk subgoaling strategy to a new pyramid subgoaling strategy formed by substituting the concept "pyramid" for the concept "disk" in rule Dsk3s.

Table 5: A Comparison of 3 Episodes			
#	state	Episode 2	Episode 3	Episode 4
1	12345,_,_	[25] Let's see... I don't think 5 will move. [26] Therefore, since 1 is the only disk I can move, and last I moved it to B, I'll put it on C this time... from A to C.	[108] I think I can do five now. [109]...Ah, it's interesting... [110] If it were five, of course, 5 will have to go to C, right? [111] So, 4 will be at B. [112] 3 will be at C. [113] 2 will be at B. [114] So 1 will go from A to C. (Fantastic.) [115] This is the way I think!!	[163] After all, it's the same thing, isn't it? [164] First, 1 will go from A to C. [165] Because, 5, at the end, will go to C, so, [166] So, 4 will go to B. [167] And then, 3 will go to C. [168] And then, 2 will go to B. [169] so, 1 will go from A to C.
2	2345,_,1	[27] so naturally, 2 will have to go from A to B.	[116] And then, 2 will go from A to B.	[170] 2 will go from A to B.
3	345,2,1	[28] And this time, too, I'll place 1 from C to B.	[117] 1 will go back from B to C. [sic]	[171] Move 1 from C to B.
4	345,12,_	[29] I'll place 3 from A to C.	[118] 3 will go from A to C.	[172] move 3 from A to C.
5	45,12,3	[30] And so I'll place 1 from B... to C. [31] Oh yeah! I have to place it on C. [32] Disk 2... no, not 2, but I placed 1 from B to C... Right? [33] Oh, I'll place 1 from B to A. (Go Ahead.) [34] Because... I want 4 on B, and if I had placed 1 on C from B, it wouldn't have been able to move.	[119] For that, um..., this time, again..., as this time 4 will have to go to B... [120] Let's move back 1 from B to A.... [121] If 4 has to go from A to B, it means... [122] 2 will have to go to 3. [123] Because 1 will... [124] So, 1 will go back from B to A.	[173] Next, 4 will go to B. So...[174] move 1 from B to A,
6	145,2,3	[35] 2 will go from B to C.	[125] And then, 2 will go from B to C.	[175] move 2 from B to C.
7	145,_,23	[36] 1 will go from A to C.	[126] And then, 1 from A to C.	[176] Move 1 from A to C,
8	45,_,123	[37] And so, B will be open, and 4 will go from A to B.	[127] And then, 4 from A to C. [sic]	[177] and then, move 4 to B.
9	5,4,123	[38] So then, this time... it's coming out pretty well... [39] 1 will... 1 will go from C... to B.	[128] And then, this time..., it's the same as before, I think..., um... [129] Of course, 5 will go to C, right? [130] For that, 3 will have to go to B, so [131] 2 will go back to A, [132] 1 from C to B.	[178] Next, 5 has to go to C, so... [179] I only need move three blocking disks to... B. [180] So, first... 1 will go from C to B,
10	5,14,23	[40] So then 2, from C, will go to... A...	[133] 2 from C to A.	[181] move 2 from C to B,
11	25,14,3	[41] And then, 1 will go from B to A.	[134] 1 from B to A.	[182] and then, move 1 from B to A.

#	state	Episode 2	Episode 3	Episode 4
		Table 5 continued: A Comparison of 3 Episodes		
12	125,4,3	[42] And then, 3 will go from C to B.	[135] 3 from C to B.	[183] Move 3 from C to B.
13	125,34,_	[43] 1 will go from A to C.	[136] 1 from A to C.	[184] 1 from A to C.
14	25,34,1	[44] What? [45] And then, 2 will go from A to B.	[137] 2 from A to B.	[185] Move 2 from A to B.
15	5,234,1	[46] And then, oh, it's getting there. [47] 1 will go from C to B.	[138] 1 from C to B.	[186] and then, 1 from C to B.
16	5,1234,_	[48] So, then, 5 will finally move from A to C.	[139] And then, finally 5 will go from A to C.	[187] And then, 5 can go to C... [188] It's easy, isn't it? [189] 5 has already gone to C. [190] Next..., 5 was able to move, because... [191] A and C were open, right?
17	_,1234,5	[49] And then, 1 will go from B to A. [50] Oh, I'll put 1 from B to C (Why?) [51] Because if 1 goes from B to A,	[140] and then, this time, um..., um, 4 will go to C, so... [141] 3 goes to A, [142] 2 goes to B [sic], [143] and then, 1 will go to A. [144] So, anyway, I will move 1 from B to A.	[192] 5 is already at C, so... [193] I will move the remaining four from B to C... [194] It's just like moving four, isn't it? [195] So... I will have to move 4 from B to C... [196] For that, the three that are on top have to go from B to A... [197] Oh, yeah, 3 goes from B to A! [198] For that, 2 has to go from B to C, [199] for that, 1 has to go from B to A. [200] So, 1 will go from B to A.
18	1,234,5	2 will go from B to C... [52] Let's try it again, ok? [53] Um... it's hard, isn't it? [54] I didn't know it would be so hard... It's hard for _me to remember... [55] And so I guess I have to do it logically and systematically.	[145] 2 from B to C.	[201] 2 goes from B to C.
19	1,34,25	[56] And 1 will go from A to C.	[146] And then, 1 from A to C.	[202] 1 will go from A to C.
20	_,34,125	[57] 3 will go from B to A.	[147] 3 from B to A.	[203] And then, 3 can go from B to A.
21	3,4,125	[58] 1 will go from C to B	[148] 1 from C to B.	[204] Then, it'll be good if 1 and 2 go to A, so. [205]...first 1 goes from from C to B,

#	state	Episode 2	Episode 3	Episode 4
			Table 5 continued: A Comparison of 3 Episodes	
22	3,14,25	[59] Because I want to move 4 to C. and to do that I have to move 2, don't I? [60] And to do that, 2 will go from C to A.	[149] And then, 2 from C to A.	[206] 2 moves from C to A.
23	23,14,5	[61] And then, 1 will go from B to A.	[150] And then, 3 from B... [151] 1 from B to A.	[207] And then, 1 moves from B to A.
24	123,4,5	[62] And then, 4 will go from B to C.	[152] And then, finally, I have succeeded in moving 4 from B	[208] Um, with this, the three at B have moved to A, so... [209] move 4 from B to C.
25	123,_,45	[63] This time, if I think of 3 on C, that will be good, so 1 will go from A to C.	[153] So, this time, um..., oh, this time, 3 naturally has to go here, so, [154] for that, 2 has to go to B. [155] so 1 will go from A to C,	[210] Next, if the three at A go to C, I will be done. [211] So first, the top two disks will be moved to B. [212] For that, 1 goes from A to C.
26	23,_,145	[64] 2 will go from A to B.	[156] place 2 from A to B,	[213] 2 goes from A to B.
27	3,2,145	[65] 1 will go from C to B.	[157] place 1 from C to B,	[214] And then, 1 goes from C to B,
28	3,12,45	[66] And then, I'll bring 3 from A to C.	[158] and then, 3 from A to C.	[215] and then, 3 goes from A to C.
29	_,12,345	[67] This time it's easy, and 1 will go from B to A.	[159] Place 1 from B to A,	[216] Oh! This time, the two on B will be moved to C. [217] Right... [218] 1 moves from B to A,
30	1,2,345	[68] 2 will go from B to C.	[160] place 2 from B to C,	[219] 2 from B to C.
31	1,_,2345	[69] And then 1 will go from A to C.	[161] and then, move 1 from A to C.	[220] And then, 1 will move from A to C.
32	_,_,12345	[70] All right, I made it.	[162] Oh yeah..., In this way, think bit by bit..., think back...	[221] I did it! [222] I think I finally got it...

References

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard.

Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*(2), 192-210.

Anderson, J.R. (1989). A theory of the origins of human knowledge. *Artificial Intelligence, 40*(1-3), 331-352.

Anderson, J. R., Farrell, R., & Saurers, R. (1984). Learning to program in LISP. *Cognitive Science, 8*, 87-129.

Anderson, J. R., Greeno, J., Kline, P. J. & Neves, D. M. (1981). Acquisition of problem solving skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Anzai, Y. (1987). Doing, understanding and learning in problem solving. In D.Klahr, P.Langley & R.Neches (Ed.), *Production System Models of Learning and Development*. Cambridge, MA: MIT Press.

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

Brown, J. S. & VanLehn, K. (1980). Repair Theory: A generative theory of bugs in procedural skills. *Cognitive Science, 4*, 379-426.

Bruner, J.S., Goodnow, J., & Austin, G.A. (1956). *A Study of Thinking*. New York: Wiley.

Chi, M.T.H., Bassok, M., Lewis, M., Reimann, P. & Glaser, R. (1989). Self explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

DeJong, G. (1988). An introduction to explanation-based learning. In Shrobe, H. (Ed.), *Exploring Artificial Intelligence*. Los Altos, CA: Morgan Kaufman.

Elio, R. & Anderson, J.R. (1981). The effects of category generalization and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning and Memory, 7*(6), 397-417.

Ernst, G.W. & Newell, A. (1969). *GPS: A Case Study in Generality and Problem Solving*. New York, NY: Academic Press.

Inhelder, B. & Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books.

Karmiloff-Smith, A. & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition, 3*, 195-212.

Karat, J. (1982). A model of problem solving with incomplete constraint knowledge. *Cognitive Psychology, 14*, 538-559.

Kowalski, B. & VanLehn, K. (1988). Inducing subject models from protocol data. In V. Patel (Eds.), *Proceedings of the Tenth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Kuhn, D. & Phelps, E. (1982). The development of problem-solving strategies. *Advances in Child Development and Behavior, 17,* 1-44.

Kulkarni, D. & Simon, H.A. (1988). The process of scientific discovery: The strategy of experimentation. *Cognitive Science, 12,* 139-175.

Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33,* 1-64.

Laird, J. E., Rosenbloom, P. S., and Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning, 1*(1), 11-46.

Langley, P. (1985). Learning to search: From weak methods to domain-specific heuristics. *Cognitive Science, 9*(2), 217-260.

Langley, P. (1987). A general theory of discrimination learning. In Klahr, D., Langley, P. & Neches, R. (Ed.), *Production System Models of Learning and Development.* Cambridge, MA: MIT Press.

Larkin, J. H. (1981). Enriching formal knowledge: A model for learning to solve problems in physics. In J. R. Anderson (Ed.), *Cognitive Skills and their acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Lawler, R. (1981). The progressive construction of mind. *Cognitive Science, 5,* 1-30.

Lenat, D.B. (1983). The role of heuristics in learning by discovery: Three case studies. In R.S. Michalski, T.M. Mitchell, & J. Carbonell (Eds.), *Machine Learning.* Palo Alto, CA: Tioga Press.

Lewis, C. (1981). Skill in algebra. In .I. R. Anderson (Ed.), *Cognitive Skills and their Acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Minton, S., Carbonell, J., Knoblock, C.A., Kuokka, D.R., Etzioni, O. & Gil, Y. (1989). Explanation-based learning: A problem-solving perspective. *Artificial Intelligence, 40,* 63-118.

Mitchell, T.M., Keller, R.M. & Kedar-Cabelli, S.T. (1986). Explanation-based generalization: A unifying view. *Machine Learning, 1*(1), 47-80.

Mitchell, T. M. , Mahadevan, S. & Steinberg, L. (1985). A learning apprentice system for VLSI design. In *Proceedings of the Third Machine Learning Workshop.* , To appear in: T.M. Mitchell, J.G. Carbonell, & R.S. Michalski (eds.) Machine Learning: A Guide to Current Research, Kluwer.

Mitchell, T. M., Utgoff, P. E. & Banerji, R. B. (1983). Learning problem-solving heuristics by experimentation. In R. S. Michalski, T. M. Mitchell & J. Carbonell (Eds.), *Machine Learning.* Palo

Alto, CA: Tioga Press.

Neches, R. (1987). Learning through incremental refinement of procedures. In D.Klahr, P.Langley & R.Neches (Ed.), *Production Systems Models of Learning and Development.* Cambridge, MA: MIT Press.

Neves, D. M. (1978). A computer program that learns algebraic procedures by examining examples and by working problems in a textbook. In *Proceedings of the second National Conference of the Canadian Society Computational Studies of Intelligence.* .

Newell, A. (in press, 19??). *Universal Theories of Cognition.* Cambridge, MA: Harvard University Press.

Newell, A. & Simon, H. A. (1972). *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S. (1980). *Competence and strategy in reasoning with common spatial concepts: A study of problem solving in a semantically rich domain.* Doctoral dissertation, Department of Psychology, University of Stockholm,

Ohlsson, S. (1987). Truth versus appropriateness: Relating declarative to procedural knowledge. In D.Klahr, P.Langley & R. Neches (Ed.), *Production System Models of Learning and Development.* Cambridge, MA: MIT Press.

Pirolli, P. & Bielaczyc, K. (1989). Empirical analyses of self-explanation and transfer in learning to program. In G. Ohlson & E. Smith (Ed.), *Proceedings of the Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Ruiz, D. & Newell, A. (1989). Tower-noticing triggers strategy change in the Tower of Hanoi: A Soar model. In G. Ohlson & E. Smith (Ed.), *Proceedings of the Cognitive Science Annual Conference.* Hillsdale, NJ: Erlbaum.

Schank, R. (1982). *Dynamic Memory: A Theory of Learning in Computers and People.* Cambridge: Cambridge University Press.

Schoenfeld, A.H., Smith, J.P. & Arcavi, A. (in press). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. In R. Glaser (Ed.), *Advances in Instructional Psychology (Vol. 4).* Hillsdale, NJ: Erlbaum.

Shrager, J. & Klahr, K. (1986). Instructionless learning about a complex device: The paradigm and obseervations. *International Journal of Man-Machine Studies, 25,* 153-189.

Siegler, R.S. (1983). Five generalizations about cognitive development. *American Psychologist, 38,* 263-277.

Siegler, R.S. (1986). Unities across domains in children's strategy choices. In Perlmutter, M. (Ed.), *Perspectives on Intellectual Development: The Minnesota Symposium on Child Psychology.* Hillsdale, NY: Erlbaum.

Siegler, R.S. & Jenkins, E.A. (1989). *How children discover new strategies.* Hillsdale, NJ: Erlbaum.

VanLehn, K. (1987). Learning one subprocedure per lesson. *Artificial Intelligence, 31*(1), 1-40.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Ed.), *Learning Issues for Intelligent Tutoring Systems.* New York, NY: Springer Verlag.

–a   VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. Posner (Ed.), *Fundamentals of Cognitive Science.* Cambridge, MA: MIT Press.

–b   VanLehn, K. (1989). Learning events in the acquisition of three skills. In G. Ohlson & E. Smith (Ed.), *Proceedings of the Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

VanLehn, K. (1990). *Mind Bugs: The origins of procedural misconceptions.* Cambridge, MA: MIT Press.

VanLehn, K. (in press   ). *Architectures for Intelligence.* Hillsdale, NJ: Erlbaum.

VanLehn, K. & Ball, W. (in press   ). Teton: An architecture for analysis of protocol data. In K. VanLehn (Ed.), *Architectures for Intelligence.* Hillsdale, NJ: Erlbaum.

VanLehn, K. & Garlick, S. (1987). Cirrus: an automated protocol analysis tool. In Langley, P. (Ed.), *Proceedings of the Fourth Machine Learning Workshop.* Los Altos: Morgan-Kaufmann.

Wallace, I., Klahr, D. & Bluff, K. (1987). A self-modifying production system model of cognitive development. in D.Klahr, P.Langley & R.Neches (Ed.), *Production System Models of Learning and Development.* Cambridge, MA: MIT Press.

Notes

[1]By "cognitive architecture," I refer to that part of human cognitive processing that is the same across all adults and all tasks. The architecture seems to encompass only fairly low level functions, such as memory, attention, perception and motor control. ACT* (Anderson, 1983) and Soar (Laird, Newell & Rosenbloom, 1987) are models of the cognitive architecture. See VanLehn (in press) for others.

[2]Anzai and Simon describe the goal-peg strategy as a kind of transition between the selective search strategy and the disk subgoaling strategy. Their description of it is rather terse and hence open to varying interpretations. For simplicity, it is ignored in the subsequent remarks.

[3]Actually, several rows were skipped, since the calculations for them were nearly identical to those for other rows. The calculations were performed on the Cascade simulation system, a stripped down version of Teton -- see VanLehn and Ball, in press.

[4]A measure of fit should weigh the accuracy of a model against its degrees of freedom or some other measure of its tailorability (Brown & VanLehn, 1980). This measure of fit is incomplete, for there is no standard measure of tailorability for rule-based models.

[5]The firings and missed opportunities that appear in the rightmost six columns of 3 could have been generated mechanically given the data in the goal utterances column, which would make interesting extension to the Cirrus automated protocol program (Kowalski & VanLehn, 1988; VanLehn & Garlick, 1987)

[6]Ruiz and Newell (1989) claim that the subject's discovery is triggered by noticing the recursive structure of pyramids, rather than the recurrent structure in the solution path. The only relevant verbal evidence is the subject's observations at line 72. Her remarks, especially the last one, "C will be used often before 5 gets there," are slightly easier to interpret under the assumption that she is talking about the solution path rather than the pyramids. However, this evidence far too weak to warrant rejection of the Ruiz and Newell hypothesis. For my analysis, it does not really matter what the subject notices. The important point is only that the noticing causes her to devise an experiment, which eventuates in discovery of the disk subgoaling strategy.

[7]The subject does not seem to form a general recursive subgoaling rule (that comes later), but instead changes her representation of the goals for the puzzle by prefixing the goal of getting disk 4 to peg B. This assumption explains why, at the three later moves that are parallel to this move (major moves 5, 21 and 29), she starts with the goal of getting 4 to B, rather than the goal of getting 5 to C as the fully-recursive strategy would do. This makes it seem quite likely that she has merely "rotely memorized" the 4-to-B goal rather than forming a general, recursive rule.

[8]The protocol was not timed, so the duration of rule acquisition events was estimated from the number of lines per acquisition event (3) and the number of seconds per line for the whole protocol (224 lines in 90 minutes, so 25 seconds per line).

[9]The above argument is a little bit circular, because the rule acquisition events were selected in

part because they showed signs of unusual cognitive activity. It is worth a moment to redo the argument without the circularity, even though that will not change the result. Suppose we define "rule acquisition events" as the first firing of a rule that has had many prior opportunities to fire. This is the definition implied by the analysis of section 3.4. With this definition, there are 3 "rule acquisition events." (The double quotes are retained as a reminder that this is a temporary definition of rule acquisition events that will be abandoned at the end of this paragraph.), and the percentage of "rule acquisition events" with extra talk becomes 66% (2/3), which is still significantly higher (p<.005) than the percentage of other moves that have extensive non-goal talk, 10% (13/127). Thus, "rule acquisition events" tend to show significantly different behavior than non-acquisition events.

[10]There is an equivalent explanation, which substitutes the notion of conflict resolution for retrieval. According to this explanation, there is no trouble retrieving the rules, but the decision about which rule to execute is influenced by the familiarity of the rules and the visual salience of the cues. The protocol data do not, of course, discriminate between the retrieval-based explanation and this explanation.

[11]As an aside, it is interesting to note that the rate of impasses is roughly the same in both subtraction and the Tower of Hanoi. According to the Sierra-based analysis of multi-column subtraction given in VanLehn (1987), only a few impasses would be needed to acquire the whole subtraction procedure -- on the order of magnitude of ten to a hundred, depending on how much context is included in the initial versions of the rules. However, the procedure's acquisition typically lasts about 50 hours (VanLehn, 1990), making the impasse rate lie somewhere in the vicinity of one impasse an hour. However, the impasses are probably not evenly distributed, because a great deal of the 50 hours is taken up with drill, which presumably increases the automaticity and retention of the procedure without changing its content. During this kind of practice, there would be no impasses. If we assume that 40 of the 50 hours are drill and only 10 hours of the subtraction curriculum is devoted to impasse-causing introduction of new material, then the impasse rate is somewhere in the vicinity of one every 10 minutes.