A Model of Acquiring Problem Solving Expertise[1]

Dennis Kibler
Rogers P. Hall

Irvine Computational Intelligence Project
Department of Information and Computer Science
University of California, Irvine 92717

INTRODUCTION

Research on computational reasoning suggests that effectively organized knowledge, not search, is the key to expert performance. Paradoxically most work in the machine learning literature assumes that learning, a form of computational reasoning, occurs in a knowledge vacuum. Concepts are acquired through search in a space of candidate conceptualizations constrained by an implicit focus provided through representational "bias." For example, in early work on learning concepts from examples, a computational learner might have been implicitly constrained by having a restricted concept description language for assimilating carefully chosen examples. Thus learning was fundamentally disconnected from background knowledge. While early efforts at computational learning provide an illuminating starting point, our work is motivated by a recognition of the importance of multiple sources of background knowledge which human learners routinely bring to novel situations. In this paper, we describe a computational model of learning problem solving skills which proceeds by tentatively connecting new situations with existing knowledge through a process of analogical reasoning. This model is in early stages of implementation.

LEARNING PROBLEM SOLVING

The nature of problem solving expertise has also interested psychological researchers. By carefully contrasting the behavior of novice and expert problem solvers, psychological research reveals that competent problem solvers construct a form of intermediate semantic representation while understanding a problem text but before attempting to generate quantitative solutions (Chi, Feltovich and Glaser, 1981; Larkin, 1983). This intermediate representation appears to consist of abstract conceptual entities (e.g., force or momentum in physics) which are related to strategic methods through problem schemata reflecting problem categories (e.g., conservation of momentum).

As a goal for a computational model of acquiring problem solving skills, expertise in problem solving can be viewed as the possession of an abstract **conceptual vocabulary** tailored to the particular problem solving domain and an assortment of **problem schemata** reflecting problem categories. These schemata provide a mechanism for retrieving appropriate problem solving strategies when a problem from a particular category is encountered. Retrieval is based on conceptual cues which must be inferred from the problem statement. Acquisition of problem solving expertise in some domain then amounts to learning these problem schemata and the conceptual vocabulary out of which they are constructed.

Solving algebra story problems

The problem domain chosen for this work is that of learning to solve "story" problems typical of algebra instruction at the elementary and high school level. For example, consider the following "triangle" problem:

> Jerry walks 1 block east along a vacant lot and then 2 blocks north to a friend's house. Phil starts at the same point and walks diagonally through the vacant lot coming out at the same point as Jerry. If Jerry walked 217 feet east and 400 feet north, how far did Phil walk?

This is called a "triangle" problem since its solution rests on relating the problem to some basic facts about triangles. Mayer (1981), analyzing a large sample of problems appearing in secondary school texts, presents a clustering of problem types consisting of approximately 50 distinct problem templates organized as a simple classification hierarchy. Levels in the hierarchy correspond roughly to families of algebraic equations underlying sets of problems and the "story lines" on which problems are based. For example, in the "amount-per-time rate" family, a variety of story templates are evident including motion, current and work problems. Although the intent of Mayer's study is to document classification schemata which students reliably use to interpret and then solve algebra story problems, identification and description of these problem schemata provide a rare opportunity for choosing a domain in which there exists a substantial body of descriptive and experimental literature based on human performance (e.g., Hinsley, Hayes and Simon, 1977; Mayer, Larkin and Kadane, 1984; Kintsch and Greeno, 1985).

[1] This work was supported by a gift from the Hughes AI Research Laboratory, a division of Hughes Aircraft.

Multiple knowledge sources

We will assume that the story problem is presented to the learning system as a relatively flat set of propositions without an explicit representation of inter-connecting structural relations. For the triangle story mentioned above this representation might be:

event(e1, walking)	*agent(e1, Jerry)*	*distance(e1, 1 block)*
	direction(e1, east)	*trajectory(e1, along lot)*
event(e2, walking)	*agent(e2, Jerry)*	*distance(e2, 2 blocks)*
	direction(e2, north)	*destination(e2, friend's)*
event(e3, walking)	*agent(e3, Phil)*	*source(e3, ?x)*
	trajectory(e3, through lot)	*destination(e3, friend's)*
	distance(e1, 217 feet)	*distance(e2, 400 feet)*

with the goal of finding *distance(e3,Answer)*. This representation is relatively complete with respect to the literal problem statement, includes information which is extraneous to the desired solution (e.g., that Jerry walks three blocks), but also excludes information which must be inferred if a solution will be found (e.g., that a right triangle has been described). Representations of similar problems described in the psychological literature (e.g., Mayer, 1981) include inferences from the data and abstractions of important details without explicitly describing how such processes take place.

In addition to choosing a representation for problem statements at the point of input, we must also describe the types of knowledge which will constitute the "multiple knowledge sources" essential for problem understanding and solution. Fortunately, the task domain of solving algebra story problems provides a finite but sizable collection of requisite sources of knowledge. On the basis of an informal content analysis of problems occurring frequently in Mayer's (1981) classification taxonomy, we will assume that a novice problem solver must have knowledge of the objects which routinely appear in problem statements, knowledge of a factual variety regarding these objects, and limited inferential abilities regarding space and time. These knowledge sources are assumed to be relatively stable during the course of learning to solve algebra story problems.

Two sources of knowledge remain, and provide the basis for what is to be learned. First, the system is to acquire a practical **conceptual vocabulary** for primitive events which are central in story problem texts. Examples are simple motion or work events, which we treat as frame-like structures that serve to organize objects described above. Second, the system is to acquire a variety of solution methods organized as **problem solving schemata**. A schema consists of a cue expressed in terms of the entities and relations of the problem description and a set of quantitative or qualitative constraints. The cue serves to facilitate recognition of situations in which the constraints are applicable. For example, matching a schema cue with a problem description in which two simple motion events occur in an opposite direction might invoke a decomposition of the problem into additive components. Event frames and problem schemata correspond to the components of problem solving expertise described in the introduction of this paper.

A PROBLEM SOLVING ARCHITECTURE

Our model of solving algebra story problems depends upon the cooperation of multiple knowledge sources operating in an asynchronous fashion through a globally accessible problem description. As shown above, the initial problem description is a jumbled set of propositions with relatively little structure. From an abstract perspective, problem solving proceeds by using existing knowledge sources to manipulate this problem description until sufficient quantitative constraints are available to allow calculation of a solution. A solution path in this space of manipulated problem descriptions consists of increasingly coherent descriptions of the current problem generated by the actions of applicable knowledge sources.

Although the problem solver may utilize a large number of knowledge sources, we can divide these various knowledge sources into four functional types. While in the previous section we described background knowledge sources primarily in terms of their content (e.g., methods, events or time), we now describe knowledge sources in terms of their role in problem solving. These include: **augmenters, organizers, asserters,** and **decomposers**. Augmenters add propositions to the problem description by applying background knowledge. In the example problem presented earlier, information about triangles would be added by augmentation. Organizers add structure to the problem description by grouping disconnected propositions into frame structures. In the example, Jerry and Phil walking from source and destination points signal that simple motion-events have occurred. Expectations encoded in the frame structure for motion-events serve to organize the problem description. Asserters contribute quantitative constraint relations that involve a single organizing frame. For example, the equation $d = r * t$ would be asserted given the cueing attributes of a motion-event, known rate and time, and unknown distance. Decomposers transform a problem into quantitatively related subproblems, yielding a constraint equation involving more than one organizing frame. For example, in a motion problem involving several driving events, the problem might be decomposed into the sum of component distances.

The problem solver schedules activities of the various knowledge sources by a competitive process which attempts to expend the least effort while still progressing towards a goal state. Augmentation and organization are least costly, but do not generate any equations, and therefore may not necessarily lead closer to a problem solution. Instead their contributions must be focused towards enabling and confirming other knowledge sources. Assertion and decomposition, although more complex, are the only operations that generate equations. They must be applied to achieve the goal. Since enabling conditions for any of the four types of knowledge sources may be only partially matched by a current problem description, we must often treat a knowledge source proposing a modification to the problem description as a hypothesis which is subject to confirmation. This is precisely the role of analogical reasoning in this model of problem solving.

Analogical Transformations in Problem Solving

In this section, we describe how recognition, elaboration and confirmation of analogies serve to connect a problem description with applicable problem schemata. The central goal of analogical reasoning is to alter the representation of the current problem so that it is partially equivalent to a previously experienced problem for which a solution method (either equations or a suitable decomposition) is known.

Analogical reasoning, then, is a process of viewing an unsolved problem *as if it were* a problem for which a solution strategy is already known. "Viewing as" in this context involves extending information from the solved problem into the description of the new problem, subject to some form of critical evaluation within the confines of the new problem. Given the importance of matching a developing problem description with known problem solving methods, analogical reasoning represents a point midway in a continuum between literal similarity and nonsensical (or anomalous) comparison. Under this assumption, reasoning and learning processes undertaken while viewing one problem as if it were another more familiar problem are comparable in kind to processes undertaken when a new problem is recognized as an instance of a known problem class.

Recognition of opportunities for transformation

Recognition of a potential analogy within our process model amounts to tentative acceptance of a problem schema on the basis of a partial match between the current problem description and a problem cue within the schema. In isolation from other processes, it is easy to imagine some form of search among existing problem schemata to find a subset of schemata which bear a promising resemblance to the target problem. The computational literature on analogical reasoning is sparse with respect to proposals for constraining such a search. Most promising is the notion of indexing potential analogs on the basis of abstract relational information (e.g., Carbonell, 1981 and Kolodner, 1983, 1984) so that recognition and retrieval will be based on higher level (and presumably more predictive) aspects of similarity. In the domain of algebra story problems, a higher level (or "systematic," according to Gentner, 1982) correspondence might be exemplified by a match between a simple motion and a simple work problem which involves some sort of event (e.g., a motion or work event), a rate and a single unknown. Such higher level descriptive aspects might be distinguished from surrounding information (e.g., the name of an agent or destination) *a priori* through supporting domain knowledge or by virtue of participating in the method associated with a problem cue in a related problem schema. Recognizing an applicable problem schema is a central activity in our model of problem solving, an activity to which processes of augmentation and organization are explicitly directed.

Elaboration and evaluation of prospective analogs

We hypothesize that recognition of analogies is identical to recognition of directly applicable problem schemata except with regard to the effort expended in confirming the match between problem descriptions. Confirmation is attained by incrementally elaborating the correspondence between aspects of a potentially applicable knowledge source and the current problem description. As aspects of the source description (e.g., organizing structural information) are extended to the problem description through elaboration, the validity of these extensions must be evaluated both in terms of the specific nature of the current problem and in terms of inferential capabilities over objects and relations in the domain of algebra story problems. In the same sense that augmentation and organization processes were described as bringing supporting knowledge sources to bear in understanding a new problem, so too can these processes serve in confirming tentative information extended from source to target problem descriptions.

For example, when confirming an analogy between a "work together" problem (two agents work on a single job together) and an "opposite direction" motion problem (agents travel in opposite directions from the same source point), knowledge of equal duration within the motion problem might be extended to the work problem by virtue of being integrally involved in the solution strategy of decomposition. The tentatively held assertion of equal duration in the work problem (extended from the motion problem) must be confirmed on the basis of knowledge which can be generated with augmentation processes (i.e., that time intervals may be additive). Similar sorts of confirmation could be achieved with organizing processes.

LEARNING MECHANISMS

The goals of learning, as described earlier, are the acquisition of problem schemata and the specialized conceptual vocabulary of which they are constructed. Broadly speaking, the learner's conceptual vocabulary (i.e., frame-like descriptions of events) gradually changes as the result of problem solving experience and instruction. Direct instruction will be involved in the extension or correction of other knowledge sources, while learning when to apply background knowledge (i.e., acquisition of problem schemata) is the learner's responsibility. Hence our computational model of learning to solve algebra story problems involves a variety of forms of learning including learning by being told, learning by taking advice and learning from examples.

The learning mechanisms we propose satisfy a number of constraints. First, they are all incremental. Thus, problem solving expertise is acquired gradually through active experience with a succession of problems and instruction presented by the teacher. Second, learning is tolerant of errors: correctness of acquired concepts or schemata is relative to problem solving experience rather than being defined in absolute. Concepts constantly change and evolve according to their problem solving utility. Third, newly acquired knowledge is connected to old knowledge by the recognition, elaboration and confirmation of analogies. Thus learning does not occur in isolation from existing knowledge of the task domain.

We now briefly consider some examples of learning processes, each example based on simple problems drawn from Mayer's (1981) taxonomy. These examples demonstrate direct interaction of instruction, inferences based on background knowledge, and problem solving experience. The first example shows how the learner refines initially naive concepts so that they are more appropriate for problem solving. The second example shows how analogy can drive the transformation of previously learned concepts and schemata into new problem solving knowledge.

Learning through instruction and experience

Before beginning, we need to describe the existing knowledge state of the learner. Suppose that the learner has no problem schemata and has a naive concept of a motion-event given by the following frame:

Motion–event: **agent:**
vehicle:
to:
from:

Assume the system is presented with the following problem:

Example 1: *Bill Less drove from Boston to Cleveland, a distance of 624 miles, in the time of 12 hours. What was his driving speed?*

Without explicit knowledge of problem types and solutions strategies (i.e., problem schemata), the learner is unable to solve any problems. Hence we allow a teacher to directly instruct the learner to use the equation $distance = rate * time$. Assuming the learner can correctly solve this problem, a number of modifications are possible for existing knowledge sources. First, the naive concept of a motion-event changes to include additional slots for distance, time and rate, since these were used in the problem solution. In addition, distance, rate and time slots are distinguished as having played an essential role in the solution process. Other slots (e.g., agent) are noted as potentially inappropriate since they did not play a role in the solution. Thus the salience of slots in an event frame with respect to achieving problem solutions is recorded. Second, the learner forms a problem schema which associates with the cue, motion-event and goal to 'find rate', the method of using the equation $distance = rate * time$.

After this episode involving direct instruction and problem solving experience, the learner's concept of a motion-event would include:

Motion-event:	**agent: (n/a)**	**distance: (essential)**
	vehicle: (n/a)	**rate: (essential)**
	to: (n/a)	**time: (essential)**
	from: (n/a)	

Note that the learner has marked the *to* and *from* slots as being not applicable (n/a) even though these may prove important during later problem solving. In fact, even if the learner deleted these slots from the representation of a motion event altogether, recovery would be possible by adding the slots again on the basis of additional experience.

Assuming further problem solving experience with simple motion problems, the system's motion-event concept might evolve to look like the following where all slots are implicitly marked "essential":

Motion-event:	**distance:**	**start-position:**
	rate:	**end-position:**
	duration:	**start-time:**
	direction:	**end-time:**

The process of adding or reinforcing slots that are used and demoting slots that are not used allows the learner to progress without an absolute notion of concept correctness. Concepts are modified in a fashion which reflects their utility. Rather than mere occurrence, inclusion and survival of components of a concept (i.e., slots) depends on their participation in the problem solving process.

Learning while reasoning analogically

We now consider a second example which demonstrates underlying processes of analogical reasoning as a guide to learning. Assume for this example that the learner is capable of solving simple distance-rate-time problems and is presented with the following problem:

Example 2: *A fisherman can catch a fish every 20 minutes. If he spends an 8-hour day fishing, how many fish will he bring home?*

We will also assume that the learner possesses a naive notion of a work-event represented as:

Work-event:	**agent:**	**pay:**
	job:	**instrument:**

The learner can recognize that this problem involves a work event, but has no problem schemata suggesting solution methods for this type of problem. In fact, it is not even possible to integrate all aspects of the problem description (e.g., the *rate* at which fish are caught) within the work event frame. However, since we are assuming that the learner has a relatively well-developed schema for simple distance-rate-time problems, it is possible to construct an analogy between this schema's cue and the current problem description. Specifically, recognition might occur on the basis of a partial match in which rates, durations and unknowns are placed in correspondence. Elaborating this analogy, work duration and rate could be associated with the naive work event frame, and a tentative correspondence between distance and work output could be formed. Thus, by viewing the simple work problem as if it were an instance of the relatively well-known class of distance-rate-time problems, the learner could attempt to solve the equation $output = rate * time$. Since this strategy yields a correct answer to the current problem, a new problem schema can be formed which associates unknown(work) and work-event with the successful equation.

Note that in this second example, the learner recognizes an analogy between problem types on the basis of a relatively remote partial matching between a schema cue and the current problem description. Although the initial kernel of correspondence is tentative, further elaboration serves to reorganize the problem description and a successful solution attempt confirms the initially tentative analogy. As a result, both the learner's conceptual vocabulary and repertoire of problem schemata are enlarged. In this example, however, recognition of the analogy between simple motion and work problems depends on a rather liberal allowance for entertaining partial matches. If the learner were allowed to attempt all such possible analogies, exhaustive search of weak partial matches over existing problem schemata cues could prove unreasonable. These problems await further study.

CONCLUSION

The preceding sections illustrate how a conceptual vocabulary and problem schemata can be learned through a combination of instruction, analogical reasoning between similar problem types, and problem solving experience. Acquisition of problem solving skills is incremental, pragmatic, and strongly connected to existing knowledge sources. Problem solving is modeled as a process of problem understanding. Through inferences supplied by processes of augmentation, organization, assertion and decomposition, problem descriptions become increasingly coherent and constrained. Analogical reasoning allows tentative application of known solution methods, subject to verification by surrounding knowledge sources.

REFERENCES

Carbonell J.G. (1981) Invariance hierarchies in metaphor interpretation. *Proceedings of the 3rd Annual Conference of the Cognitive Science Society*, 292–295.

Chi M.T.H., Feltovich P.J and Glaser R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science 5*, 121–152.

Gentner D. (1982) Are scientific analogies metaphors? In D. Miall (Ed.) *Metaphor: Problems and Perspectives*. New Jersey: Humanities Press.

Hinsley, Hayes and Simon (1977) From words to equations: meaning and representation in algebra word problems. In P.A. Just and M.A. Carpenter (Eds.) *Cognitive Processes in Comprehension*. Hillsdale, New Jersey: Lawrence Erlbaum and Associates, 89–106.

Kintsch W. and Greeno J.G. (1985) Understanding and solving word arithmetic problems. *Psychological Review 92*(1), 109–129.

Kolodner J.L. (1983) Maintaining organization in a dynamic long-term memory. *Cognitive Science 7*, 243–280.

Kolodner J.L. (1984) Towards an understanding of the role of experience in the evolution from novice to expert. GIT-ICS-84/07.

Larkin J.H. (1983) The role of problem representation in physics. In D. Gentner and A. Stephens (Eds.) *Mental Models*. LEA: Hillsdale, N.J., 75–98.

Mayer R.E. (1981) Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science 10*, 135–175.

Mayer R.E., Larkin J.H. and Kadane J.B. (1984) A cognitive analysis of mathematical problem-solving ability. In R.J. Sternberg (ed.) *Advances in the Psychology of Human Intelligence*. LEA: Hillsdale, New Jersey, 231–273.