

Case-Based Learning: Predictive Features in Indexing

COLLEEN M. SEIFERT

Department of Psychology, University of Michigan, 330 Packard Road, Ann Arbor, MI 48104

KRISTIAN J. HAMMOND

Department of Computer Science, The University of Chicago, 1100 East 58th Street, Chicago, IL 60637

HOLLYN M. JOHNSON

Department of Psychology, University of Michigan

TIMOTHY M. CONVERSE AND THOMAS F. McDOUGAL

Department of Computer Science, The University of Chicago

SCOTT W. VANDERSTOEP

Department of Psychology, University of Michigan

Editor: Michael Pazzani

Abstract. Interest in psychological experimentation from the Artificial Intelligence community often takes the form of rigorous post-hoc evaluation of completed computer models. Through an example of our own collaborative research, we advocate a different view of how psychology and AI may be mutually relevant, and propose an integrated approach to the study of learning in humans and machines. We begin with the problem of learning appropriate indices for storing and retrieving information from memory. From a planning task perspective, the most useful indices may be those that predict potential problems and access relevant plans in memory, improving the planner's ability to predict and avoid planning failures. This "predictive features" hypothesis is then supported as a psychological claim, with results showing that such features offer an advantage in terms of the selectivity of reminding because they more distinctively characterize planning situations where differing plans are appropriate.

We present a specific case-based model of plan execution, *RUNNER*, along with its indices for recognizing when to select particular plans—appropriateness conditions—and how these predictive indices serve to enhance learning. We then discuss how this predictive features claim as implemented in the *RUNNER* model is then tested in a second set of psychological studies. The results show that learning appropriateness conditions results in greater success in recognizing when a past plan is in fact relevant in current processing, and produces more reliable recall of the related information. This form of collaboration has resulted in a unique integration of computational and empirical efforts to create a model of case-based learning.

Keywords: case-based reasoning, indexing, modeling, planning, analogical reasoning

1. Learning in humans and machines

In many machine models of learning, the goal seems to be the *mimicking* of human learning at the level of input and output behavior. However, I/O matching is problematic because human and machine behaviors are fundamentally different in their environments and background knowledge. Even in novel domains, it is nearly impossible to achieve access to large quantities of background knowledge (as humans do with problem-solving strategies, analogical reminders, prior processing contexts, and episodic traces of similar events) in

machine learners. How can humans and machines be performing the same processing given such different knowledge bases? One possible fix to this comparison problem is to equate background knowledge by choosing tasks where little or no knowledge is involved. This strategy was used in the development of EPAM (Feigenbaum, 1990) as a computer model of human verbal learning, where the input stimuli (nonsense syllables) are designed to provide little connection to past knowledge. However, EPAM involved a part of human behavior that is, by experimental design, far outside of the normal learning circumstances that an individual encounters in the world. When do people learn in the absence of *any* knowledge?

Another alternative is to attempt to provide the machine with background knowledge for some task comparable to that of human subjects. When the domain is very restricted (*e.g.*, inflating balloons (Pazzani, 1991)), this approach may be successful. However, in most domains (for example, in commonsense planning), there is no limited set of rules that can be considered sufficient. Without comparable knowledge bases, a machine learning model should be dissimilar to a human learner for the same reasons that a water plant near the Great Lakes is unlike one at the edge of the Gobi desert: they must produce the same product under extremely different operating circumstances. We face great difficulty in trying to “test” machine and human learners under similar conditions of environment, experience, and background knowledge.

Therefore, we argue for the importance of testing not I/O performance, but *paradigms*. What needs “testing” is the underlying assumptions of any machine learning model—the specific processing claims of its approach that should obtain across other tasks and learning situations, and that distinguish the approach from other competing theories. Newell (1991) argues for this approach, suggesting that one should first determine whether one has the basics of the class of models that will work before worrying about exactly matching detailed behavior. By examining the assumptions of a computational model, we can determine its psychological plausibility while avoiding the temptation of repeatedly adjusting it in order to exactly match human behavior in a specific test task.

In the following paper, we present an example of this approach to models of learning. We start with a problem: how does a case-based learner find relevant information in memory? We begin with a hypothesis that the indices with predictive utility in the planning domain are those structural features that distinctively indicate when particular plans are appropriate. A series of studies with human learners is then presented, contributing an explanation of what role predictive features play in retrieval. We then examine one case-based learning model in particular, RUNNER, and describe the implementation of the predictive features hypothesis for plan execution called “appropriateness conditions.” Finally, we present experiments designed to examine whether appropriateness conditions are effective in human learning.

2. Predictive features hypothesis: What indices are learned?

Previous research has demonstrated the utility of learning by analogy (Carbonell, 1983; Gick & Holyoak, 1983; Gick & Holyoak, 1980; Hammond, 1989; Pirolli & Anderson, 1985; Ross, 1989). However, one must first *retrieve* a relevant candidate case from memory. Given that memory is full of past experiences, only a small number of which may be

relevant, successful access requires identifying a past case with important similarities to the current situation, while distinguishing among other cases with similar features. How experiences are encoded into memory, and what types of cues may provide access to them, is called the *indexing* problem. How to find a relevant analog in memory is a central issue in research on analogical learning in psychology (Anderson, 1986; Gentner, 1983; Gentner & Landers, 1985; Gick & Holyoak, 1980; Holyoak, 1985; Pirolli & Anderson, 1985; Rattermann & Gentner, 1987; Ross, 1987; Ross, 1989; Seifert, McKoon, Abelson, & Ratcliff, 1986). While most computational approaches have ignored the problem of retrieving past experiences, one approach—case-based learning (Hammond, 1989; Kolodner, 1985; Riesbeck & Schank, 1989)—has attempted to determine the type and source of features used to index cases in episodic memory during learning.

2.1. *Types of indices: Surface vs. structural*

By most psychological accounts, retrieval depends on how similar the new problem is to a target example stored in memory, given a context of other related competing cases (Anderson, 1986; Gentner & Landers, 1985; Tversky, 1977). However, previous psychological research on indices has focused on a single factor in explaining when access occurs: the abstractness of the features (Gentner, 1983; Holyoak, 1985). Specifically, features that are more superficially (surface) related to the intended analogical meaning are contrasted with features that involve more deep, thematic (structural) relations among pattern elements (Gentner, 1983; Ross, 1987). Gentner (Gentner & Landers, 1985; Rattermann & Gentner, 1987) has found that surface features result in more frequent access in a memory retrieval task, and Ross (1987) has demonstrated that different kinds of surface similarity (story line versus object correspondence) lead to different rates of access. One possible explanation for these results is that the ability to make use of structural features is limited even when they are available; however, this does not appear to be the case. In these same studies, more abstract, relational features also reliably produced access to past cases based on structural features alone (Gentner & Landers, 1985; Rattermann & Gentner, 1987; Ross, 1987). Other studies have also shown activation of prior cases in memory based solely on abstract, thematic cues, particularly if subjects are instructed to attend to them (Seifert et al, 1986).

Why, then, do superficial features appear to promote better case access compared to abstract features? Surface features may be more readily available (require less inference) than structural features, playing a role in memory access before any abstract features are even available (Hammond & Seifert, 1992). Surface features may also serve to identify prior examples when individual cases are unique in content within the memory set (as in Rattermann and Gentner, 1987, where there was only one base “squirrel” story in memory). Unfortunately, in many real-world domains, there is substantial overlap of surface features between cases, so that abstract features are particularly important. For example, access based on structural features alone may be necessary when learning in a new domain, where past experiences won’t share many surface features with new problems. In these cross-contextual reminders (Schank, 1982), abstract strategies from one domain are applied in another, such as taking the “fork” strategy from chess and developing it as the

“option” play in football (Collins, 1987). Access to past information based on abstract features is particularly important when the common features are incomplete or ambiguous (Seifert, 1992).

The capacity to learn a general principle in one setting and transfer it to other, non-superficially related settings, is the essence of intelligent behavior. Even if people are only rarely able to make use of structural features, it is critical that we determine *when* such transfer does occur, and how we might accomplish this process in machine learners.

2.2. *Useful indices: Causal vs. correlated*

Our research on planning suggests that structural indices—those representing the abstract relationships among goals and plans—serve to constrain plan choices in a given situation (Hammond, 1989; Hammond & Seifert, 1992). Because this causal information serves to distinctly identify types of planning problems, the types of solution strategies that can be applied, and potential failures to avoid, it also provides useful indices to past plans. Thus, our claim is that among all possible features in a planning situation, only a limited set of these features—those that are relevant to the way in which the current causal interaction of goals and plans can be changed—are predictive of planning constraints, and therefore most useful as indices.

Two examples illustrate features most useful in predicting when plan knowledge is appropriate:

Flight of the Phoenix

X was working late on a project that was due in a matter of days. As he saw the deadline approach, he considered the following two plans: either continue to work straight through the night (and the next day) or get a good night’s sleep and come back to the office refreshed. The first plan allows the use of all of the time for work on the project. The second provides less time, but the time it does provide is of a better quality.

While thinking about his problem, X was reminded of a scene from the movie “Flight of the Phoenix” in which a character played by Jimmy Stewart had to start a damaged plane in the desert. The plane’s ignition needed explosive cartridges, and Stewart had only seven left. The plane’s exhaust tubes were also filled with sand, which had to be blasted out using the same explosive cartridges. Stewart’s character was faced with a choice: either try to start the plane using all the cartridges directly, or use some of them to clear the exhaust tubes to enhance the overall utility of the other cartridges. Stewart decided to use some of the cartridges to clear the lines, thus optimizing the likelihood of success for the remaining cartridges in starting the engine.

This situation involves a specific type of resource conflict, and the reminding carries with it information about how to make the decision. In the next example, the reminding is much more concretely related to the task at hand:

Missed Exit

Y was driving along an expressway in the left lane, because traffic was moving faster there. He spotted the exit where he wanted to get off. Unfortunately, by the time he worked his way into the right-hand lane, he had missed the exit.

The next time Y was driving on the same expressway in the left lane, intending to get off at the same exit, he again remembered the exit too late to make the exit, but noticed it had been preceded by a large billboard advertising “Webber’s Inn.”

The third time, Y noticed the billboard, remembered the exit, and got into the right lane in time to make the exit.

In the Flight of the Phoenix example, the movie situation shares no surface features with the write-or-sleep decision, but on the structural level the two situations match exactly: Each protagonist has a goal and a limited resource and must decide between two plans for using the resource; one plan uses the resource directly in service of the goal, while the other splits the resource between an optimization step and a direct-use step. These features suggest the solution of focusing attention on the question of whether sleeping would leave enough time to complete the project.

In the Missed Exit example, the features used in indexing are much more concretely related to the task at hand. Rather than predicting a causal relationship, the billboard simply serves as a correlated feature that is easier to notice. Structurally, the problem is to optimize over two goals: drive fast (hence drive in the left lane), and get off at the right place; however, a structure-level “solution” would look like “Move into the right-hand lane only when you need to,” which does not provide detectable features. By recalling the earlier failures, Y is able to modify the features used to retrieve the exit goal from memory so as to make optimal use of observable surface features that are correlated with the presence of the exit. Through experience, non-causally related features that predict the desired choice point can also be learned.

These two examples represent opposite ends of the causal spectrum—in one, the features are predictive because they are causally related; in the other, the features are predictive because they happen to be correlated with the desired event. Thus, predictiveness is not an inherent feature of the indices themselves, such as “surface” or “structural” properties, but of the *relative* role played in characterizing episodes within a domain (Seifert, 1988). In some situations, surface features alone may be sufficient to predict what past knowledge should be accessed. For example, in learning several rules of statistics, retrieving past examples based on surface features like “the smaller set being mapped onto the larger one” may be sufficient for the learner’s purposes (Ross, 1989). In other situations, structural features may be the *only constants* available across a variety of planning situations. Together, these examples present challenges that any model of learning must be capable of answering.

The point of these examples is that indexing of information will occur based on any and all predictive features within the task context. Structural feature reminders may be more likely to carry information needed to solve complex problems, while surface-level reminders may be more likely to help us react quickly to our environment. In both cases, the indexing features can be expected to consist of those that reliably predict the utility of

the retrieved information. This suggests that indexing should be based on any features that are apparent to the planner at the time when retrieval would be useful.

3. Predictive features in human learning

In human learning, is memory access possible through a subset of structural features that help to select plans? From a functional perspective, prior cases must be indexed using features that will be readily apparent to the processor at the time when retrieval is desired. In planning situations, this point may be defined as when conditions and constraints are known, but a decision has not yet been made. That way, the retrieved plan can suggest possible solutions or warn of potential dangers while the situation outcome is yet undetermined. Features that predict when prior cases might be useful could be better retrieval cues because they more specifically describe relevant planning constraints in the current processing context. If human memory is indeed operating under the predictive features hypothesis, features related to *when* plans are relevant should lead to better access to past experiences in memory than other, equally associated features.

For example, consider this story:

A chemist was trying to create a new compound designed to allow preservation of dairy products stored at room temperature. The chemist was so confident that his experiments would succeed that he went ahead and ordered several truckloads of fresh dairy products to be delivered to demonstrate the utility of the new compound.

This story contains features with *predictive utility*; that is, the features present in the story allow the retrieval of past experience and predict a possible planning failure *before* the complete structural analog (namely, *counting your chickens before they're hatched*) is even available.

In order to examine the issue of predictive features in indexing, we conducted several experiments (as reported in Johnson & Seifert, 1992) comparing the retrieval of prior cases based on structural cues. The critical question is, are some subsets (cues predictive of planning failures) better than other (equally similar) cue subsets in retrieving prior cases? A set of common structural features was determined within narrative stories based on common cultural adages such as "closing the barn door after the horse is gone" (Seifert, McKoon, Abelson, & Ratcliff, 1986). These themes are based on knowledge about problems in planning that can occur, and how to avoid or solve them (Lehnert, 1980; Dyer, 1983), and so are also likely to be familiar patterns to subjects.

According to the predictive features hypothesis, structural features involving planning errors should form a privileged subset that leads to more successful case access than would be attained using other sets of features. Other shared structural features should be less useful as indices. To test this claim, we used a reminding paradigm based on Gentner and Landers (1985). Previous studies established that this retrieval task paradigm produces results similar to problem-solving tasks (Ross, 1989). In the current experiments, retrieval alone served as the dependent measure, but the stories all involved planning content. In a single experimental session, subjects were given a set of base stories to read. After a ten minute distractor task, subjects were given a set of cue stories and asked to write down any

Table 1. Example feature sets for the Chemist story.

Theme: Counting your chickens before they're hatched.

Base story elements:

- 1) X desires A.
- 2) X assumes A.
- 3) X does not act to ensure A.
- 4) X invests resources based on A.
- 5) A does not occur.
- 6) X has wasted resources.

Story elements contained in each cue story type:

Complete: 1-6

Predictive: 1-3

Outcome: 4-6

base stories that came to mind. The cues used in the experiments were of three types: a complete-theme cue, a predict-theme cue, and an outcome-theme cue. Table 1 shows the abstract features for one of the themes used, and which of those features each cue story type contained.

The critical decision point precedes the action of going ahead and investing resources based on anticipated success. It is *before* taking this action that one can make a choice of pursuing a different plan. Once made, one has committed to "counting the chickens," whether or not they hatch as planned. Therefore, the failure itself is not considered a predictive feature, since its presence already determines the set of outcomes. The predictive features set contained thematically-matching information up to the point of a planning decision; the other, conclusion features set included the thematically matching information from the point of decision through the outcome of the story. Examples of the stories used in the experiments are given in Table 2.

Each of the cue stories was followed by instructions taken from Gentner and Landers (1985): "If this story reminds you of a story from the first part of the experiment, please write out the matching story as completely as you can. Try to include the names of characters, their motives, and what happened." The responses were scored based on whether they matched the intended base story, matched a different study story, matched no study stories, or was left blank. The rate of reminders was compared to chance estimates, one of which measured the probability of access given demonstrated availability in free recall for each subject.

The results showed that, while both subsets of cues resulted in reliable retrieval of study stories, the cue set that included features predictive of a planning decision proved to contain better cues than the set including the planning decisions themselves and their outcomes. The main findings were that there were no differences in the number of *matches* for predictive and outcome cues. However, only predictive cues were matched at a higher-than-chance level because the predict-theme cues led to significantly fewer responses involving *mismatched* stories than did outcome features. Therefore, the stories containing predictive features led to more *reliable* access to matching stories in memory.

In follow-up experiments, we asked subjects to directly compare each type of partial cue to the base stories used in the experiments to determine whether the cues were differentially distinguishing the base stories, based on a) an overall similarity judgment not biased towards

Table 2. Example Study and Cue Stories in the Experiments.

Sample study story:

Judy was overjoyed about the fact that she was pregnant. She looked forward to having a baby boy, and wanted one so badly she felt absolutely certain it would be male. As a result, she bought all kinds of toy cars, trucks, miniature army soldiers, and even arranged an extravagant "It's a boy" party. Finally, the big moment came, and she was rushed to the hospital. Everything went smoothly in the delivery room, and at last she knew. Judy's lively bouncing baby was actually a girl.

Complete-theme test cue:

Harrison disliked his small apartment and shabby furniture. His rich aunt Agatha was near death, and although he hadn't seen or spoken to her in 15 years, he felt assured of inheriting a great fortune very shortly because he was her only living relative. He had already thought of plenty of ways to spend a lot of money fixing his place up. Confident of his inheritance, Harrison began charging everything from color televisions to cars to gourmet groceries. When Aunt Agatha finally died and her will was read, she had left all her millions to the butler and now Harrison was in debt.

Predict-theme test cue:

Harrison disliked his small apartment and shabby furniture. His rich aunt Agatha was near death, and although he hadn't seen or spoken to her in 15 years, he felt assured of inheriting a great fortune very shortly because he was her only living relative. He had already thought of plenty of ways to spend a lot of money fixing his place up.

Outcome-theme test cue:

Confident of his inheritance, Harrison began charging everything from color televisions to cars to gourmet groceries. When Aunt Agatha finally died and her will was read, she had left all her millions to the butler and now Harrison was in debt.

themes and b) thematic similarity (using thematic similarity instructions for the rating task from Gentner and Landers (1985)). The results showed that subjects rated the sets of pre-decision (predictive) cues and the sets of decision and post-decision (outcome) cues as equally similar to the base stories, whether using just thematic or overall similarity as a standard; thus, the reminding results are not due to any differences in length of cue or amount of information in the two types of cues.

In a final experiment, we asked subjects to match pre-decision stories and post-decision stories to the base stories, as in the second experiment, except that the subjects *directly* compared the stories rather than using the partial cue stories to retrieve base stories from memory. The results showed that the pre-decision cues produced more reliable matches than the decision and post-decision cues. Predictive features thus appear to more *distinctively* characterize relevant planning situations, leading to reminders appropriately specific to the planning decision, and few other reminders. The predictive features better characterize individual themes as distinct from other themes. While both the predict-theme and the theme-outcome cues were found to provide reliable access to matching cases, the predictive features showed an advantage in terms of the *selectivity* of reminding. This indicates that the elements in the predict-theme stories distinguished the themes more clearly, and thus subjects tended either to find the right story or to retrieve nothing. The theme-outcomes, however, tended to evoke a wider range of intrusion responses, indicating that the features available in them were also shared by other potentially retrievable episodes. Causal features involving goal and plan interactions will, when encoded, distinguish among situations where different plans will be appropriately applied.

4. Predictive features in case-based models

In the following section, we describe current research on a specific model of indexing in case-based learning, RUNNER, which incorporates this notion of the predictive features needed for successful indexing. Like other case-based systems (Alterman, 1985; Hammond, 1989; Kolodner, 1985; Kolodner & Simpson, 1989; Simpson, 1985; Riesbeck & Schank, 1989), RUNNER makes the following basic claims:

- New solutions are built from old solutions.
- Intelligent learning arises from anticipation of difficulties and optimizations based on prior experience.
- Prior experiences are selected from memory through matching indices.
- Useful indices in the planning domain involve features of goals, features associated with past successes, and features predicting problems to avoid.

A distinctive thrust of this case-based approach to learning (as opposed to other types of AI systems) is the incorporation of “lessons learned” into memory so that past errors can be retrieved and avoided appropriately. Case-based learners accomplish this through a specific strategy: *anticipate and avoid* problems. To do this, case-based learners keep track of the features in their domains that are *predictive* of particular problems, so those problems can be anticipated in future situations where they are likely to arise. As a result, the use of predictive features in indexing allows a case-based learner to find relevant past experiences in memory, thereby providing the chance to bootstrap from experience. Thus, learning is supported by the ability to store and access *actual* past problems, rather than attempting to anticipate through all *possible* ones (as in search models such as SOAR (Newell, 1991)).

In case-based learning, experiences encoded into the memory base reflect the structure of a domain (*e.g.*, more examples of the goals that tend to arise in conjunction, the interactions between steps that tend to occur, and successful solutions) and are built up incrementally through learning. This approach involves the study of the actual *content* of features used in the organization of memory. In particular, we are interested in the type of features that may be predictive within a particular knowledge domain and task. A processor would benefit from retrieving related past cases at any time that they can provide information that is helpful in the current situation.

4.1. Predictive features in planning: Appropriateness conditions

Our argument has been that it is not possible to characterize the utility of indices on the basis of inherent type (*e.g.* surface vs. structural). Rather, the types of indices that turn out to be predictive depend strongly on the domain, the task, and the type of processing normally performed for the task.

The RUNNER project focuses on questions of retrieval in the context of routine plan execution. In this task, predictive features are those indices that signal that a particular plan

may be used with success, and should be attempted. The problem of plan use is somewhat different than the more general problem of planning—in particular, retrieval of standard plans in routine situations can often be more directly tied to perceptually available features, and this is advantageous if it can increase the efficiency of a frequently-used plan.

It might seem that the indices appropriate for plan retrieval are simply the preconditions of the plan, since a plan can be executed when its preconditions are satisfied, and not otherwise. However, some preconditions are not at all predictive of plan relevance; for example, “having money” is a precondition for many plans, yet having this precondition met does not necessarily predict a prime occasion for executing any of these plans. Instead, it would be useful to learn those features that should be explicitly considered when deciding whether to embark on a given plan. These *appropriateness conditions* provide memory access to the plan itself when circumstances are favorable to plan execution.

The plan then serves both as a memory organization point for annotations about the current progress and problems of the use of the plan, and as a hook on which to hang past experiences of its use. Its appropriateness conditions determine whether it is relevant to selecting current actions to be executed.

4.2. *Appropriateness conditions in* RUNNER

Learning in many domains can be characterized as the acquisition and refinement of a library of plans. “Plan” refers to the collection of explicitly represented knowledge that is specifically relevant to repeated satisfaction of a given set of goals, and which influences action only when a decision has been made to use it. The plans in the library that are used will be incrementally debugged and optimized for the sets of conjunctive goals that typically recur (Hammond, 1989). The RUNNER project centers around plan use in a commonsense domain. By examining case-based learning within a planning task, we have uncovered a specific indexing vocabulary related to the control of action in the world.

The representational system for RUNNER’s memory, as well as the bulk of the algorithm for marker-passing and activation, is based on Charles Martin’s work on the DMAP parser (see (Martin, 1990)). The memory of RUNNER’s agent is encoded in semantic nets representing its plans, goals, and current beliefs. Each node in RUNNER’s plan net has associated with it a (disjunctive) set of *concept sequences*, which are a (conjunctive) listing of states that should be detected before that plan node can be suggested.

In the plan network, this amounts to a nonlinear hierarchical plan net, where the activation of plan steps (that are not ordered with respect to each other) can be dependent on environmental cues that indicate a particularly good time to perform that step.

Nodes in the plan net become activated in the following ways:

- “Top-level” (the basic level instantiating different action sequences) plans become activated when the goal they serve is activated, and a concept sequence indicating appropriate conditions is completed.
- Specializations of plans are activated by receiving a *permission marker* from the abstract plan, in addition to the activation of a concept sequence.

- Parts (or steps) of plans are also activated by completion of a concept sequence, and by receiving a permission marker from their parent.

Once activated, many plans have early explicit verification steps which check if other conditions necessary for success are fulfilled, and abort the plan if not. Passing of permission markers is not recursive, so that the state information indicating an opportunity to perform a sub-plan must be recognized for execution to proceed further. This means that individual sub-plans must have associated with them concept sequences that indicate opportunities to be performed. (For a more complete explication, see (Hammond, Converse, & Martin, 1990)).

As an example, RUNNERS plan for making coffee involves a number of steps, many of which are independent in terms of ordering. Among these is a step for taking a filter from a box of filters and installing it in the coffeemaker. When the coffee plan becomes active, this step and others receive permission markers from the plan, which means that they can now be activated if they “see” one of their concept sequences completed. One of the sequences for the filter-installing step is simply the conjunction of seeing the filter box and being close to it. This activates the plan to install the filter. After doing one perceptual check (that there are indeed filters in the box), the step is performed. This style of sub-plan activation has the advantage that opportunities to perform steps can be taken without having explicitly planned the opportunity.

Appropriateness conditions in RUNNER, then, are the union of the concept sequences of plan nodes and any initial verification steps required. Recognizing situations under which it is appropriate to invoke a plan is handled by testing the set of preconditions for the plan. Under assumptions of perfect knowledge and a closed world, there is little divergence between preconditions and the appropriateness conditions. When these assumptions are relaxed, however, there are several different ways in which the divergence can become important in plan execution and reuse:

- A precondition can effectively be “always true”. This means that the plan may depend upon it for correctness, but an executor will never run into trouble by not worrying about its truth value. This sort of fact should not be an “appropriateness condition” since consideration of it cannot help in the decision whether to use the plan.
- A precondition may be almost always true, and it may be difficult to know or check in advance. If the consequences of an abortive attempt at performing the plan are not too severe, then this sort of fact should not be an appropriateness condition since the utility of knowing its truth is outweighed by the cost of acquiring the knowledge.
- A precondition may be intermittently true, but may be easily “subgoaled on” in execution, and achieved if false. (This of course depends strongly on representation of plans, and how flexible the execution is.) To the extent this can be handled in “execution”, the condition should not be an appropriateness condition since whether or not the condition holds, the plan is likely to succeed.
- A particular condition may not be a precondition *per se*, but may be evidence that the plan will be particularly easy to perform, or will produce results that are preferable to the

usual default plan for the relevant goals. This *should* be an appropriateness condition, even though it is not a precondition.

The power of appropriateness conditions is that they help to select the set of features that an agent should consider before deciding on a given plan.¹ This predictive features set is interestingly different from both the (possibly infinite) set of facts that need to be true for the plan to work, and also the set of facts explicitly used in the plan's construction. Even if an agent is assumed to have immutable plans in memory that will determine its competence, there is still room for learning the particular appropriateness conditions that govern when to invoke particular plans. The set of predictive features for indexing in RUNNER, then, is the set of appropriateness conditions needed to identify *when* a potential plan may be relevant.

4.3. *The predictive utility of appropriateness conditions*

Predictive yet only probabilistic conditions are included as appropriateness conditions even though they are not true preconditions because they are a concern of the agent in deciding on the viability of a plan. Here, as with the indexing of plans and repairs (Hammond, 1991), actions need to be indexed by the conditions that favor their applicability. As new conditions are learned and others are found to be unreliable, this set of predictive features changes.

Part of the process of refining appropriateness conditions can be taken care of by relatively simple "recategorization" of various conditions in the taxonomy sketched above, in response to both failure and unexpected success. Here are some ways in which this sort of recategorization can be applied:

- Drop appropriateness conditions that turn out to be always true. At its simplest, this is merely a matter of keeping statistics on verification steps at the beginning of plans.
- If a plan fails because some sub-plan of it fails, and that sub-plan failed because some appropriateness condition didn't hold, then *promote* that condition to the status of an appropriateness condition for the superordinate plan. That is, make the use of the larger plan contingent on finding the condition to be true.
- If a plan is frequently found to have false appropriateness conditions in situations where the plan is needed, and the conditions are under the agent's control, consider including the conditions in an *enforcement* plan that maintains them, so that the conditions can then be assumed true for that plan.

At present, the RUNNER program reliably handles the first two types of learning listed above, and both of them are entirely empirical. That is, RUNNER starts to omit particular sensory verification steps when they turn out to be always true, and, when possible, RUNNER verifies conditions that have turned out to determine failure of a plan before undertaking the body of the plan. Continuing research on the problem of recognizing opportunities for enforcing conditions will focus on prioritizing learning for goals considered most important.

A major part of learning to plan effectively, then, is the development of a library of conjunctive goal plans, and the simultaneous tuning of the plans, their appropriateness

conditions, and the environment itself to maximize the effectiveness of the plans. This approach requires:

1. having a plan library with optimized plans for the different sets of goals that typically recur;
2. for each plan, using indices consisting of appropriateness conditions, which are easily detectable and indicate the conditions under which the plan is appropriate;
3. *enforcing* standard preconditions so that they can be assumed true.

Thus, in RUNNER, predictive features are determined by separating preconditions that define the plan structure from the conditions that signal *appropriate use* of the plans. This separation allows tuning the conditions under which an agent will consider a given plan independent of its defined structure. This learning process improves the match between an environment and use and reuse of a library of plans in memory. Vocabularies for indexing (the appropriateness conditions in RUNNER) are thus designed to index actions in terms of the circumstances that favor their utility.

This type of utility assessment is distinguished from the type of consideration done in systems such as PRODIGY (Minton, 1988) in two ways. First, the assessment done in RUNNER concerns the utility of the features used to control the access of plans rather than the utility of the plans themselves. This allows RUNNER to actually store and reuse plans that might not be useful *in general* but are highly applicable in specific situations. Second, RUNNER's assessment is incremental and does not depend on the existence of a complete domain model. As such, it also allows the system to both remove and promote features to *appropriateness* status.

5. Indexing plans in human learning

Predictive features were found to show an advantage in terms of the selectivity of reminding, producing more reliable recall of related information. This predictive features hypothesis was then implemented in the case-based model RUNNER. These "appropriateness conditions" identified structural planning features that indicate when a relevant plan should be selected. Because of the variation in surface content in planning problems, the features one must recognize to access past plans will often be ones that relate to the causal structure of the planning situation. These appropriateness conditions should include the same features a planner needs to detect and monitor anyway during the planning process.

Two experiments were conducted to explore the effect of explicitly teaching the appropriateness conditions for two plans within a complex lesson (see Vanderstoep & Seifert, 1992). Rather than using specific examples from the RUNNER project, which require a lot of interaction with the environment during planning, we chose stimuli that had the same critical properties: a variety of features could be used to index the plans in memory; a set of structural features are present; and a subset of these features could be identified as "appropriateness conditions" that predict when a particular plan is appropriate for execution. Our domain involved elementary probability theory, which has been shown to produce

Table 3. Examples of the two similar probability problems.

Permutations problem: An art contest has 10 artists who have entered a painting. Two awards will be given: one to the first-place artist and one to the second-place artist. How many different ways can the artists win the awards?

Combinations problem: An art contest gives away three awards each year for "Outstanding Young Artist." All of the awards are identical—a \$100 cash prize. There are 15 artists who are eligible for this award. How many different ways can the eligible artists win the three awards?

learning of a variety of indexing features to retrieve relevant probability principles (Ross, 1989). These prior studies demonstrated that human learners can and do attend to both structural features and surface features in order to select a relevant principle. However, the paradigm also provides a testbed for the role of appropriateness conditions in plan selection: Will the presence of structural features predicting when a plan may be used facilitate human learning?

In these experiments, subjects studied either a similar pair of principles (combinations and permutations) or a dissimilar pair of principles (combinations and conditional probability) through example problems and formulas. Examples of the similar principles are given in Table 3.

The primary manipulation was whether or not learners received information regarding the appropriateness conditions that indicate *when* to apply each principle. For these plans, the appropriateness conditions were defined as whether the order of the set is considered when counting the number of possibilities. Half of the subjects received appropriateness conditions information that identified what aspects of the problem (whether the order of the objects in the sets being counted was important) determined when each of the principles should be used. The other half of the subjects received no instructions on appropriateness conditions, but were asked to review the study information for an equal time interval. It is hypothesized that when people are learning principles that are very similar to each other (i.e., easily confused), knowledge of the appropriateness conditions of the principles will be very important for learning to apply the principles. Will instructing learners about the appropriateness conditions of these problem-solving principles improve later performance?

The answers to the test problems were scored for correctness using a system similar to Ross (1989). The two groups showed no differences in the number of problems completed accurately. However, subjects instructed about the appropriateness conditions made fewer confusion errors than subjects in the review-only condition whenever a similar problem pair was used; however, this instructional manipulation had no effect when a dissimilar problem pair was used.

Why did the appropriateness conditions group make fewer errors than the review-only group? Consider that performance on this problem-solving task is a function of 1) recognizing the appropriate principle, and 2) remembering the specific content of the procedure (e.g., in a mathematical problem-solving task, the particular variables and how they are arranged in the formula), and 3) using the principle correctly in the new problem. Because the appropriateness conditions subjects made fewer errors, but they did not solve more problems correctly, the instructions may have improved just the ability to identify *when* to apply a procedure, but may not improve memory for the formula, or the ability to implement it.

A second experiment examined this notion by testing subjects' *selection* of a principle for a new problem separate from their ability to execute the plan correctly. Subjects learned two of three principles of probability theory using the same materials as in the first experiment. However, instead of asking subjects to solve the test problems (as in Experiment 1), in this experiment they were told simply to select the formula (by name) among four response choices: the names of the two principles the subjects studied, along with the formulas, "neither formula" and "don't know (can't remember)." Finally, they were asked to provide an explanation for their answer.

The subjects informed about the appropriateness conditions did in fact perform better at the more specific task of selecting when to use each formula. No difference was found between the appropriateness conditions group and the review-only condition for low-similarity problems. However, subjects informed about appropriateness conditions did better than subjects in the review-only condition at selecting the correct planning principle to use. The results also suggest that subjects informed about the appropriateness conditions provided significantly better explanations for their answers than subjects in the review-only condition. In both the high-similarity and low-similarity pairs, no differences were found between the review-only and the appropriateness conditions groups for memory of any of the formulas.

These studies support the notion that instructions on appropriateness conditions facilitate the *selection* of appropriate solution procedures in solving probability problems. With the appropriateness conditions instructions, subjects received information about the importance of the order of the objects, how to identify when order is or is not important, and when each principle should be applied. When learners were provided with this information, they were more likely to correctly detect situations when each formula should be used. These experiments show that when plans are similar, providing appropriateness conditions instructions does improve selection performance while not affecting memory for the formula or the ability to instantiate the formula correctly in a target problem.

Although it may seem intuitively obvious that teaching appropriateness conditions would be helpful, instruction is not often designed like this. Consider how these same principles might be presented in a probability textbook. Permutations might be taught in one section, followed by combinations taught in the next section. Students could become proficient at solving each of the two different types of problems; however, without explicit instruction on what makes a certain problem a permutations problem and what makes another problem a combinations problem, learners may have inadequate knowledge of when each formula should be applied.

These results confirm the predictive features hypothesis, and illustrate the need for appropriateness conditions when learning in a domain with similar plans. Much of the previous work on memory retrieval of analogies and cases has focused on the features that determine similarity; however, these results suggest distinctiveness, as well as similarity, is very important for recognizing when to apply prior knowledge.

6. Implications for indexing in learning models

Our computational and empirical results have shown that plans are indexed in memory based on the features that predict their applicability. These results confirm the predictive features

hypothesis as developed in the RUNNER model, and serve as strong verification of the case-based learning approach as a psychological model. When indexing cases in memory, appropriateness conditions—indices related to when to consider a particular plan—are more useful than equally related information that is not helpful in distinguishing among plan options. Thus, from a functional perspective, the most useful features would be those that let one predict and access potential problems and solutions whenever they are relevant to current processing. This definition of predictive features constrains the set of indices one might propose for a computational model using past experiences, and also serves as a hypothesis for the features expected to be evident in human learning.

Past approaches have suggested that the level of abstraction of indexing features, in terms of the surface versus structural dichotomy, predicts what features will be helpful in learning by analogy (Gentner & Landers, 1985; Rattermann & Gentner, 1987; Ross, 1987). However, the evidence from the experiments using structural features presented here, and from the performance based on structural features in RUNNER, suggests that all structural features are not the same; instead, the distinctiveness—the ability to select among related plans—of a subset of structural features, and not the overall similarity or abstractness, determines valuable indices in memory.

Specifying predictive features in a task domain requires establishing the causality inherent in the domain for characterizing plan examples, and identifying both surface and structural features that are predictive of important decisions within the task context. By using a causal analysis of the goal interactions as indices for storing planning information in memory, it is possible to access plan strategies applicable to the problem (Hammond & Seifert, 1992). These strategies provide the planner with alteration techniques and information as to what parts of the initial causal configuration are appropriate targets of change, leading to specific plans for the current situation. This type of causal feature vocabulary has been incorporated into models that design tools for programs (Birnbaum & Collins, 1988), plan in the cooking domain (Hammond, 1989; Kolodner, 1987), plan radiation treatments (Berger & Hammond, 1991), schedule deliveries (Hammond, Marks, & Converse, 1988), run errands (Hammond, Converse, & Martin, 1990), and learn geometry (McDougal & Hammond, 1992).

In some domains, knowledge of the causal factors in goal and plan interactions will be needed in order to characterize when to use particular plans. For example, natural categories reflect learning the features “animate” and “inanimate” as critical concepts (Smith & Medin, 1981; Mandler, 1991). This is not simply because it is *possible* to divide the natural world into living and nonliving things; certainly, many other such criteria are possible (such as, external versus internal gestation) and are correlated with desirable information. Rather, recognizing the class of objects that are animate, and therefore capable of self-initiation, allows one to make plans that take into account possible actions on the part of those objects. The “animate” feature serves to activate expectations about how these actions may affect one’s own goals. Thus, it is an important predictive feature in the natural world when making decisions about how to deal with objects in the environment. Within planning domains, then, we propose that the commonalities in the features that serve as useful indices are likely to be based on those features that predict successful pursuit of one’s goals.

Other, equally related surface and structural features may be present in a domain, such as preconditions necessary for plan execution. However, unless these features are capable

of distinguishing one goal pursuit situation from another, learning them is in a sense “academic”: only distinctive features, whether surface or structural, provide predictive indices. A purely correlational learner such as Anderson’s rational analysis model (Anderson, 1991) could find some predictive features, if the world is kind enough to isolate them in different types of causal problems. For example, correlated features may be used to signal problem types (as “incline plane” problems in physics) (Chi, 1988). While this approach is often successful, it will be unable to learn the types of non-observable structural features predictive of goal satisfaction in *RUNNER*. Unless these causal features are computed and encoded into memory, a model will be unable to learn the important features needed to distinguish appropriate plans across problem contexts. Of course, correlational models could be adapted to incorporate this ability to attend to structural and predictive features; however, exactly how to accomplish this, and a theory of what those features are likely to be in the planning domain, is what our predictive features model provides.

7. Conclusion

Our approach to learning predictive features not only describes what people *do* learn, it also suggests what people *should* learn in a functional sense, and therefore what knowledge learning programs must be capable of acquiring. In this collaborative enterprise, the psychological experiments confirm that human memory access to prior cases can be facilitated by cues that contain particular predictive structural features. This evidence lends support to case-based models as potential models of human reasoning and memory; more specifically, our results have shown that human behavior is consistent with the case-based learning tenant that cases are indexed by the features that predict their applicability. The empirical studies in this collaboration confirm the *ideas* behind our computational models, rather than the specific *behaviors* of a particular implementation.

Acknowledgments

This work was supported by the Office of Naval Research under contract N0014-85-K-010 to the University of Chicago and contract N0014-91-J-1128 to the University of Michigan. Support was also provided by the Defense Advanced Research Projects Agency under contract F49620-88-C-0058 to the University of Chicago, DARPA contract number N00014-91-J-4092 monitored by the Office of Naval Research, and Office of Naval Research grant number N00014-91-J-1185. Thanks to Michael Pazzani and anonymous reviewers for comments on an earlier manuscript.

Notes

1. One alternative approach is to design the agent so that it considers only relevant conditions, either by hand-crafting the decision procedure, or by a mixture of hand-crafting and clever compilation of declarative specifications (Rosenschein, 1986).

References

- Alterman, R. (1985). Adaptive planning: refitting old plans to new situations. *Cognitive Science*.
- Anderson, J.R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum.
- Berger, J., and Hammond, K. (1990). ROENTGEN: A case-based approach to radiation therapy planning. In *The Proceedings of the 32nd Annual Scientific Meeting of the American Society of Therapeutic Radiology and Oncology*.
- Berger, J., and Hammond, K.J. (1991). ROENTGEN: A memory-based approach to radiation therapy treatment design. In (R. Bareiss, Ed.), *Proceedings of the 1991 DARPA Workshop on Case-Based Reasoning*. Los Altos, CA: Morgan Kaufmann, Associates.
- Birnbaum, L., and Collins, G. (1988). The transfer of the experience across planning domains through the acquisition of abstract strategies. In J. Kolodner (Ed.), *Proceedings of a Workshop on Case-based Reasoning*. Los Altos, CA: Morgan Kauffmann, p. 61–79.
- Carbonell, J.G. (1983). Learning by analogy: Formulating and generalizing plans from past experience. In R.S. Michalski, J.G. Carbonell, and T.M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach*. Los Altos, CA: Morgan Kaufman.
- Dyer, M.G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Cambridge, MA: MIT Press.
- Feigenbaum, E. (1990). The simulation of verbal learning behavior. In J. Shavlik, and T. Dietterick (Eds.), *Readings in Machine Learning*. New York: Morgan Kauffmann.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.
- Gentner, D., and Landers, R. (1985). Analogical reminding: A good match is hard to find. In *Proceedings of the International Conference on Systems, Man, and Cybernetics*. Tucson, AZ.
- Gick, M.L., and Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306–355.
- Gick, M.L., and Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Hammond, K.J., Marks, M., and Converse, T. (1992). Planning, opportunism and memory. In D. Atkinson, and D. Miller (Eds.), *Planning Systems for Autonomous Mobile Robots*.
- Hammond, K.J., Converse, T., and Martin, C. (1990). Integrating planning and acting in a case-based framework. In *The Proceedings of the 1990 National Conference of Artificial Intelligence*.
- Hammond, K. (1989). *Case-based planning: Viewing planning as a memory task*. San Diego: Academic Press, Inc.
- Hammond, K.J., and Seifert, C.M. (1992). Indexing plans in memory. *Proceedings of the Fourteenth Annual Cognitive Science Society Conference*, Bloomington, IN.
- Hammond, K.J., and Seifert, C.M. (1993) A cognitive science approach to case-based planning. In S. Chipman and A.L. Meyrowitz (Eds.), *Foundations of Knowledge Acquisition: Cognitive Models of Complex Learning*. Norwell, MA: Kluwer Academic Publishers, 245–267.
- Hammond, K.J., Seifert, C.M., and Gray, K.C. (1991). Functionality in analogical transfer: A hard match is good to find. *Journal of the Learning Sciences*, 1:2, 111–152.
- Johnson, H.M., and Seifert, C.M. (1992). The role of predictive features in retrieving analogical cases. *Journal of Memory and Language*, 31, 648–667.
- Johnson, H.K., and Seifert, C.M. (1990). Predictive utility in case-based memory retrieval. In *Proceedings of the Twelfth Annual Cognitive Science Society*, Boston, Massachusetts.
- Kolodner, J.L., and Simpson, R.L. (1989). The mediator: Analysis of an early case-based problem. *Cognitive Science*.
- Kolodner, J.L. (1984). *Retrieval and organizational strategies in conceptual memory: A computer model*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Kolodner, J.L., Simpson, R.L., and Sycara, K. (1985). A process model of case-based reasoning in problem-solving. In *The Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, Los Angeles, CA: IJCAI.
- Kolodner, J. (1987). Extending Problem Solver Capabilities Through Case-Based Inference. In *Proceedings of the Fourth International Workshop on Machine Learning*. Los Altos, CA: Morgan Kaufman Publishers, Inc., p. 167–178.
- Kolodner, J.L. (1983). Reconstructive memory: A computer model. *Cognitive Science*, 7, 281–328.
- Lehnert, W. (1980). Plot units and narrative summarization. *Cognitive Science*, 5, 293–331.
- Martin, C.E. (1990). Direct Memory Access Parsing. Ph.D. Thesis, Yale University.
- Martin, C.E., and Reisbeck, C. (1986). Uniform parsing and inferencing for learning. In *Proceedings of the Fifth Annual Conference on Artificial Intelligence*, Philadelphia, PA: AAAI.
- McDougal, T., Hammond, K.J., and Seifert, C.M. (1991). A functional perspective on reminding. *Proceedings of the Thirteenth Annual Cognitive Science Society*, Chicago, Illinois.
- McDougal, T. and Hammond, K. (1992). A recognition model of geometry theorem-proving. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Bloomington, Indiana.
- Minton, S. (1988). Learning Effective Search-Control Knowledge: An Explanation-Based Approach. Technical Report # 133, Carnegie-Mellon University, Department of Computer Science.
- Newell, A. (1991). *Unified theories of cognition*. Cambridge: Harvard University Press.
- Owens, C. (1990). Functional Criteria for Indices and Labels. *Working Notes of the AAAI Spring Symposium Series*. Palo Alto, CA.
- Pazzani, M.J. (1989). Indexing Strategies for goal specific retrieval of cases. *Proceedings of the Case-Based Reasoning Workshop*, Pensacola Beach, FL.
- Pazzani, M.J. (1991). *Learning Causal Relationships: An integration of empirical and explanation-based learning methods*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pirolli, P.L., and Anderson, J.R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology*, 39, 240–272.
- Ratterman, M.J., and Gentner, D. (1987). Analogy and similarity: Determinants of accessibility and inferential soundness. In J. Anderson (Ed.), *Proceedings of the Ninth Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ; Lawrence Erlbaum Associates, Inc.
- Riesbeck, C.K., and Schank, R.C. (1989). *Inside Case-based Reasoning*. Hillsdale, NJ: Erlbaum.
- Rosenschein, S.J., and Kaelbling, L.P. (1986). The Synthesis of Digital Machines with Provable Epistemic Properties. In *Proceedings of 1986 Conference on Theoretical Aspects of Reasoning About Knowledge*.
- Ross, R.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629–639.
- Ross, B.H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456–468.
- Ross, B.H. (1989). Reminders in learning and instruction. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning*. Cambridge: Cambridge University Press.
- Schank, R.C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. New York: Cambridge University Press.
- Schank, R.C., and Abelson, R.P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schank, R.C., and Riesbeck, C. (1990). *Inside case-based reasoning*. Hillsdale, NJ: Erlbaum.
- Seifert, C.M. (1988). Goals in Reminding. *Proceedings of the DARPA Work on Case-based Reasoning in AI*, Clearwater Beach, Florida.

Seifert, C.M. (1993) "The role of goals in analogical retrieval." K.J. Holyoak and J.A. Barnden (Eds), *Advances in Connectionist and Neural Computation Theory, Vol. 2: Analogical Connections*. Norwood, NJ: Ablex Publishing Company.

Seifert, C.M. McKoon, G., Abelson, R.P., and Ratcliff, R. (1986). Memory connections between thematically similar episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 220–231.

Simpson, R.L. (1985). A Computer Model of Case-based Reasoning in Problem-solving: An Investigation in the Domain of Dispute Mediation. Ph.D. Thesis, School of Information and Computer Science, Georgia Institute of Technology.

Smith, E.E., and Medin, D.L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

VanderStoep, S.W., and Seifert, C.M. (1993). "Learning 'how' vs. learning 'when:' Improving problem solving transformer." *Journal of the Learning Sciences*.

Vanderstoep, S., and Seifert, C.M. (1993). "Learning WHEN: Predictive features in analogical transfer." P. Pintrich, D. Brown, C. Weinstein (Eds.), *Perspectives on Student Motivation, Cognition, and Learning: Essays in Honor of Wilbert J. McKeachie*. Hillsdale, NJ: Erlbaum Associates.

Received March 31, 1992

Final Manuscript January 18, 1994