Improving Case Retrieval Through Observing Expert Problem Solving

Michael Redmond*
College of Computing
Georgia Institute of Technology
Atlanta, Georgia 30332-0280
E-mail: redmond@cc.gatech.edu

Abstract

As case-based reasoners gain experience in a domain, they need to improve their case retrieval so that more useful cases are retrieved. One problem in doing this is that the reasoner who most needs to learn is least able to explain successes or failures. A second problem is that uncontrolled pursuit of an explanation could be very expensive. There are three keys to the approach presented. First, the student observes expert problem solving and sets up expectations for what the expert will do next. When expectations fail, the reasoner has its failure isolated to a single step, and the correct action for the situation has been provided. Second, if the student can retrieve part of a case that would have suggested a correct prediction, then that case snippet can be used to limit the explanation process, making the process more efficient. Third, when no explanation can be found, the reasoner resorts to empirical adjustment of feature importance.

Case-based reasoning (CBR) is based on the observation that experience, retained in the form of cases, can be used to efficiently and effectively solve future, similar problems. A case-based reasoner can improve in a number of different ways. It can acquire new cases. Or it can improve its case retrieval, so that more useful cases are retrieved. In Redmond (1989b) we discussed our approach to acquiring new cases through apprenticeship. Part of apprenticeship involves observing and understanding expert problem solving. This same kind of experience can be used to improve retrieval of cases.

Improving case retrieval is one of the key issues in case-based reasoning. Novice reasoners are frequently most influenced by surface features in retrieving previous experiences (Ratterman and Gentner 1987; Ross 1987). In becoming more expert, a reasoner must learn to retrieve *usefully* similar cases.

Much of the work on improving case retrieval has focused on learning indices though explaining a reasoner's own successful or unsuccessful problem solving. Learning indices based on explanations requires at least three things:

1. Realizing the need to learn.
2. Determining what the correct result should be.
3. Assigning credit for successes or blame for failures.

The problem is that these can be hard. A novice reasoner is the *most* in need of improvement and the *least* prepared to learn. A novice may not be able to generate the correct result. A novice may also have trouble assigning credit or blame. How can a novice get around these problems?

Apprenticeship can provide assistance with this problem. A novice can observe an expert solving a problem. A good student, who actively follows along with the example, sets up expectations at each point in the expert's problem solving. When the expectations are incorrect then he has a failure. The student thus has immediate feedback. The student realizes the need to learn. The student has been given the correct result. Most important, the failure has been isolated to a single step. This tighter feedback loop enables learning when it otherwise might not be possible, and makes learning more efficient.

An example will illustrate some of the issues involved in improving case retrieval through observing an expert. A student is observing an instructor solve an automobile diagnosis problem. This is part of an ongoing mentor relationship. The instructor has checked whether the car stalls when cold. The instructor has adjusted the idle mixture screw and determined that the car still stalls. The instructor has tightened any loose spark plugs, and determined that the car still stalls. He has checked if the throttle dashpot is out of place (it wasn't), and if the fuel level in the carburetor was too high (it was). Figure 1 shows the effect of all this on the problem solving context and also shows some of the more general features of the problem that were elicited by the instructor. At this point the reader need only note that there are a

*This research has been supported by the Army Research Institute for the Behavioral and Social Sciences under Contract No. MDA-903-86-C-173, and Contract No. MDA-903-90-K-0112 and by DARPA contract F49620-88-C-0058 monitored by AFOSR. The author wishes to thank Janet Kolodner for her advice and guidance, and Ashwin Ram for helpful comments on earlier versions of the paper.

Current Context
Complaint: Car Stalls Out
Other Symptoms: None Reported
Frequency of Problem: intermittent
Temperature When Fail: Cold
Type of Car: 1980 Chrysler LeBaron
Car Owner: Julie Crider
Mechanics Involved: Tom Davis
Ruled In:
Spark Plug Connections Ok
Idle Mixture Ok
Carburetor Fuel Level High ***
Ruled Out:
Low Idle Speed
Lean Idle Mixture
Incorrect Position Throttle Dashpot
Tests Done:
Car Stalls When Cold?
Spark Plugs Loose?
Car Still Stalls?
Too Small a Distance Between Throttle Dashpot Stem and Throttle Lever?
Too High Level of Fuel in Carburetor Float Bowl?
Test Results:
Car Stalls When Warm
All Loose Spark Plug Connections Tightened
Car Still Stalls
Distance Between Throttle Dashpot Stem and Throttle Lever = 2cm
Distance Between Carburetor Float and Choke Chamber Surface = 3cm
Fixes Done:
Increase Position Idle Mixture Screw
All Loose Spark Plug Connections Tightened

Figure 1: Part of Current Context.

number of contextual features, including complex values for things that had been ruled in and ruled out, tests done etc.

At this point, the student has set up an expectation using a part of a case from his episodic memory. The student had previously been a participant or an observer of that step. Figure 2 shows most of the context at the time of the predicted action's previous occurrence. For short, we call this the *'incorrect' context*. The student's experience suggests the hypothesis that the choke linkage is sticking. We call this the *'incorrect' prediction* to indicate that it is not an appropriate prediction for the current situation. The instructor makes a different hypothesis, that the carburetor needle valve leaks. The student could have made this prediction, which we call the *'correct' prediction*. In the past he had observed the instructor taking that action. The context from that previous time is shown in Figure 3. We will call that the *'correct' context*. Why is the instructor's action a better choice for the current problem solving? The incorrect context matches a good number of the features of the current problem. The key difference favors the instructor's action, however. The information in the current context that the carburetor fuel level is high is the key difference. This information favors the correct hypothesis that the carburetor needle valve leaks. How can the student improve his case retrieval so that he will make the correct prediction in the future?

Incorrect Context
Complaint: Car Stalls Out
Other Symptoms: None Reported
Frequency of Problem: intermittent
Temperature When Fail: Cold
Type of Car: 1979 Chrysler Cordova
Car Owner: Bill Moss
Mechanics Involved: Tom Davis, Kevin Cousins
Ruled In:
Spark Plug Connections Ok
Idle Mixture Ok
Idle Speed Ok
Ruled Out:
Lean Idle Mixture
Low Idle Speed
Incorrect Position Throttle Dashpot
Idle System Leak Air
Tests Done:
Car Stalls When Cold?
Spark Plugs Loose?
Car Still Stalls?
Too Small a Distance Between Throttle Dashpot Stem and Throttle Lever?
Idle System Leak Air?
Test Results:
Car Stalls When Cold
All Loose Spark Plug Connections Tightened
Car Still Stalls
Distance Between Throttle Dashpot Stem and Throttle Lever = 2cm
No Apparent Air Leaks
Fixes Done:
Increase Position Idle Mixture Screw
Increase Position Idle Speed Screw
All Loose Spark Plug Connections Tightened

Figure 2: Portions of incorrect context.

Correct Context
Complaint: Car Stalls Out
Other Symptoms: None Reported
Frequency of Problem: daily
Temperature When Fail: Cold
Type of Car: 1981 Chrysler Cordova
Car Owner: Paul Crider
Mechanics Involved: Kevin Cousins
Ruled In:
Spark Plug Connections Ok
Idle Mixture Ok
Carburetor Fuel Level High ***
Idle Speed Ok
Ruled Out:
Lean Idle Mixture
Idle System Leak Air
Tests Done:
Car Stalls When Cold?
Spark Plugs Loose?
Car Still Stalls?
Idle System Leak Air?
Test Results:
Car Stalls When Cold
All Loose Spark Plug Connections Tightened
Car Still Stalls
No Apparent Air Leaks
Fixes Done:
Increase Position Idle Speed Screw
All Loose Spark Plug Connections Tightened

Figure 3: Portions of correct context.

When the student's expectations are not met, then the student realizes the need to learn. Apprenticeship also takes care of the need for the student to know the correct action. The action taken by the instructor is assumed to be correct. The student still must determine what features of the current situation make the correct prediction appropriate, and/or make the incorrect prediction inappropriate. The student's exposure to that action in a previous example helps. He knows the context in which the action was previously taken. There are a number of similarities and differences among the contexts. A purely empirical approach will place some of the credit or blame on some spurious features that don't make a difference. On the other hand, if the student tries to analyze all of the feature values in the contexts and explain why one prediction is more relevant than the other, that will be expensive. Also, we certainly can't assume that a student will always be able to explain why one prediction is more appropriate. How can these opposing forces be reconciled?

Our solution makes use of a combination of analytical and empirical methods. Similarity-based methods are used to focus explanation. When the student is able to explain the appropriateness or inappropriateness of a prediction, the associated case part is marked with that indication. The indication can be positive, that the prediction made by that part of the case is appropriate in a situation, in which case we call it an index. The indication can also be negative, that the prediction is not appropriate in a situation. We call the latter a *censor*. In addition, the student empirically adjusts the importance of matching different features. We first present the analytical approach to learning indices and censors. Then we discuss the empirical adjustment.

Analytical Approach - Learning Indices and Censors

The purpose of learning indices and censors is to improve prediction and diagnosis when they are carried out through CBR. Redmond (1990a) discussed our case representation. Briefly, cases are divided into *snippets*, each of which contain the information relevant to the pursuit of one primitive goal. Predictions are generated by retrieving a case snippet from memory. Each snippet contains the context in which it occurred. This enables similarity assessment, as well as the comparisons between correct and incorrect contexts mentioned above. An index is a particularly salient set of features of the context in which the snippet occurred. During retrieval, if a situation matches some of the indices, then that significantly increases the possibility that that snippet will be retrieved to provide guidance. The indices can be *parts* of contextual features. For example, *one* of the test results found in the problem solving leading up to the snippet might be marked as an index. A censor is a state (part of a feature) that suggests that a case snippet is not appropriate in that situation. If in a future situation that state exists, then the snippet can be rejected during retrieval.

We call our method *analytical feature comparison*. It has five steps which involve distinguishing responsible features through first comparing feature values, then trying to explain the relevance of differences. The initial feature comparisons are similarity-based, the analysis comes into play in latter steps. The opportunity to apply the process occurs when the student, observing the instructor, uses a case snippet to incorrectly predict the instructor's action. The process starts out by retrieving a snippet that would have predicted the instructor's

Current Context Ruled In:
Spark Plug Connections Ok Idle Mixture Ok
Carburetor Fuel Level High
(a): Part of the current problem solving context.

Correct Snippet (case-generate-hypoth-305) Context
Ruled In:
Spark Plug Connections Ok +++ Idle Mixture Ok +++
Carburetor Fuel Level High +++ Idle Speed Ok
(b): Part of the correct snippet's problem solving context.

Incorrect Snippet (case-generate-hypoth-312) Context
Ruled In:
Spark Plug Connections Ok +++ Idle Mixture Ok +++
Idle Speed Ok
(c): Part of the incorrect snippet's problem solving context.

Figure 4: Portions of snippets involved in learning an index.

action. For short, we will call this snippet the *'correct' snippet*. The action itself is an effective additional cue that enables the correct snippet to be retrieved now. We will call the snippet that the student used to make the incorrect prediction the *'incorrect' snippet*. After the correct snippet has been retrieved, the process is ready to begin.

1. Eliminate from consideration features which are the same in both snippets' contexts. Features compared include aspects of the diagnosis, such as the tests done and their results, things ruled out, etc., as well as domain dependent features, such as car type. Given the contexts in Figures 2 and 3, the complaint, the other symptoms, and the ambient temperature when the failure occurs are the same in both snippets. Therefore, they don't provide a way of distinguishing why the correct snippet is appropriate in the current context. These are therefore no longer considered candidates for explanation.

2. Compare the remaining features with the current problem solving context. Features in which the current context better matches the correct context than the incorrect context are selected as candidates. It is more likely that something in those features would indicate that the correct snippet is appropriate. For example, in Figure 1 the current context's 'ruled in' feature had the value shown in Figure 4(a). The correct context and the incorrect context had the 'ruled in' values shown in Figure 4(b) and 4(c). Matches to the current context are shown by '+++'s. The incorrect context's 'ruled in' was not as close a match as the correct snippet's 'ruled in'. Thus 'ruled in' will continue to be included in the candidates. At the same time, the incorrect snippet's 'fixes done' is a better match to the current context than the correct snippet's 'fixes done'. Therefore, 'fixes done' will be eliminated from consideration for explanation. In the example, this step eliminates the frequency of the problem, the items 'ruled out', and the mechanics involved, since they favor retrieval of the incorrect piece.

```
Correct Snippet (case-generate-hypothesis-305) Context
  Type of Car:          1981 Chrysler Cordova
  Car Owner:            Paul Crider ***
  Ruled In:
      Spark Plug Connections Ok
      Idle Mixture Ok
      Idle Speed Ok
      Carburetor Fuel Level High ***
  Tests Done:
      Car Stalls When Cold?
      Spark Plugs Loose?
      Car Still Stalls?
      Idle System Leak Air?
      Too High Level of Fuel in Carburetor Float Bowl? ***
  Test Results:
      Car Stalls When Cold
      All Loose Spark Plug Connections Tightened
      Car Still Stalls
      No Apparent Air Leaks
      Distance Between Carburetor Float and
      Choke Chamber Surface = 2cm ***
```

Figure 5: Remaining portions of correct snippet after second step of analytical feature comparison.

3. Isolate the parts of the correct snippet's feature values that cause the remaining features to better match the current context. This is to narrow the responsibility to *parts* of features which favor the retrieval of the correct snippet in the current context. These parts of features are the best candidates for distinguishing why the correct snippet is appropriate in the current context. In the current example, for items 'ruled in', the isolated part is that the carburetor fuel level is high. This is the clause that makes the correct snippet's 'ruled in' better match the current 'ruled in'. Figure 5 shows the features remaining after the second step of the process. The parts of features isolated by this third step are marked with '***'s. This is where the explanation process will be focussed. This completes the initial feature comparison, which limits the analytical reasoning done.

4. Try to explain the significance of each of the remaining parts of features. The learner can try to relate each to the current action taken by the instructor. The student in the example tries to explain the relationship between the 'Ruled In' that the carburetor fuel level is high and the hypothesis that the carburetor needle valve leaks. An explanation that the student could (depending of his level of knowledge) construct is that a leaking carburetor needle valve could enable fuel to keep flowing into the carburetor float bowl, thus causing the fuel level to become high.

For learning censors a greater variety of relationships are useful. The learner can try to relate each remaining feature part to the current action taken by the instructor, or to the action suggested by the incorrect snippet. The relationships themselves can be different. A part of the current context relating to a 'contradiction' of the incorrect action is a good indication of a need for a censor. For example, if in a second problem *Carburetor Fuel Level Normal* has been 'ruled in', that contradicts an incorrect snippet's hypothesis that the carburetor needle valve leaks.

5. If such a relationship can be found, then the partial feature value is marked as an index or censor. In the main example, *Ruled In: Carburetor Fuel Level High* can be saved as an index to the correct snippet. In the same or similar situation in the future the snippet will then be more likely to be retrieved as a source of predictions or diagnostic actions.

In the second example, the incorrect snippet can be annotated with the censor *Ruled In: Carburetor Fuel Level Normal.* In the same or similar situation in the future the snippet will not be retrieved as a source of predictions or diagnostic actions.

We should note some limitations of our approach to learning indices and censors. First, the matching of parts of features has to be exact. Some form of knowledge-based matching like that used by Bareiss (1989) and Koton (1988) is needed. Fortunately, substituting knowledge-based pattern matching for simple matching doesn't affect the method. It just substitutes better reasoning capability in steps 1 through 3. Second, there is no attempt to learn indices that involve a conjunction of feature parts. There are certainly situations in which such an index would be necessary. With better explanatory capability, the student could learn more sophisticated indices. Some of the limitations on explanation may need to be relaxed though, costing the process some efficiency. Both of these limitations will be addressed in future work.

Empirical Feature Comparison

Sometimes a learner may not be able to explain why a prediction is wrong, or why another one might be better. The learner is not an expert. When he cannot, he needs to resort to less powerful, less knowledge-intensive methods. He can attempt an empirical approach to improving case retrieval. As with learning indices and censors, apprenticeship helps identify the need to learn. It also isolates the problem to a single step, and provides the correct action for the situation. As with learning indices and censors, the learner can attempt to retrieve part of another case that would have suggested a correct prediction. If such a case snippet can be retrieved, then the contexts can be compared. The differences can be used to empirically adjust the weights on features in the matching function. The adjustment method places greater importance on features that match and lead to correct predictions. It places less importance on features that match and lead to incorrect predictions. We should note that there are results that suggest that people can learn to distribute their attention among features, giving different weights to each (Nosofsky 1987). We originally discussed this approach in (Redmond 1989a). We have now integrated it with the analytical approach discussed in the previous section.

Feature	'Incorrect'	'Correct'	Change in Importance
CAR-TYPE	Partial	No Match	less important
CAR-OWNER	Match	No Match	less important
COMPLAINT	Match	Match	no change
HOW-LONG	Match	Partial	less important
RULED-OUT	Partial	Match	more important
TESTS-DONE	Partial	Match	more important
FIXES-DONE	Partial	Match	more important
CURRENT-HYP	Match	Match	no change
WHEN	Partial	Slight	less important

Figure 6: Example empirical saliency adjustment.

Figure 6 shows how the empirical change is done on an example incorrect prediction. Those features of the current context that match the context of the correct snippet more closely than the context of the incorrect snippet are made more important. Those features of the current context that match the context of the incorrect snippet more closely than the context of the correct piece are made less important.

The student can also make empirical adjustments when he is successful. When the student correctly predicts the instructor's action, the features of the current problem solving context that matched the features in the previous case are made slightly more important. In future problem solving these empirical adjustments lead to the features considered important by the instructor having more influence on the case pieces retrieved.

Related Work

We have addressed the issue of making case retrieval better. We use a form of apprenticeship to isolate a failure, and to obtain the correct solution step. Retrieval of other previous case snippets and comparison of features is used to focus explanation. Through explanation, indices and censors are learned.

Other approaches have been suggested for learning indices and censors. Hammond (1986), and Simpson (1985)'s approaches explained instances of their own successful problem solving. This requires a significantly more expert reasoner because the correct steps must be generated. Hammond and Simpson also emphasized creating indices to avoid failures. However, in trying to explain failures their approaches had to consider all features. Our approach limits the consideration of features to a smaller set. In addition, our approach avoids having to match the situation to an abstraction, as in Hammond's approach, an expensive proposition. Therefore, our approach saves explanatory effort.

In Barletta and Mark's (1988) Explanation-Based Indexing approach, the reasoner attempts to explain the relationships of features of the case to the actions taken. Those features for which an explanation can be formed are made indices of the case. Our approach takes more information into account in order to limit explanatory effort. Thus our approach is more efficient.

Bareiss (1989)'s PROTOS created both censors that specified when a case was not appropriate, and 'difference links' which specified the condition in which one solution should be avoided and another used. But the instructor did all the distinguishing of which features should be considered in explaining the failure. PROTOS doesn't include the initial consideration of an extra case that our approach does. So our approach offers the benefit of increased efficiency of learning through initial similarity-based comparison to other cases.

At a high level, our approach shares similarity with Lebowitz (1986)'s suggestion to use similarity based measures to focus explanation. However, we have applied an instantiation of that general idea to learning indices and censors.

Empirical adjustment of feature importance was also suggested by Aha (1989) and Salzberg (1988) in the context of purely empirical instance-based learning. The contribution here is its use in conjunction with analytical approaches, as a fallback strategy when the reasoner doesn't possess the knowledge necessary to distinguish why an action is appropriate or inappropriate.

Evaluation

Through the process of observing an expert, a reasoner can significantly improve its ability to predict the expert's actions, and thus its ability to diagnose. The improvement comes through acquisition of new cases, learning indices and censors, and through adjusting feature salience. Our system, CELIA, which is based on the model, has been evaluated through presentation of a sequence of 24 examples of expert problem solving. Twenty random orders of the examples were presented. The performance measure was the average accuracy of the system's predictions of the expert's actions. The improvement over the course of exposure to examples is seen in the data presented in Figure 7.

An ablation study showed that the by the end of the example sequences, the effect of the learning methods is starting to be seen. We lesioned the functions which learn indices and censors, and which adjust feature saliency. By the last 6 examples in the sequences, the average performance advantage for including these approaches was two percent. This difference is not that large, but for some problems the advantage was as high as 10 percent.

We expect that the effects will be larger with greater experience. To this experience level, not that many indices are learned. [1] Many of these are learned late in the sequences, leaving little time for them to be useful. The system doesn't have a strong domain model, so in many instances it cannot explain the importance of different feature values. With greater experience, this problem will gradually become less of a factor. In addition,

[1] An average of 1.2 are learned per observed example. An average example has 37 steps, which become 37 snippets. The range is from 18 steps to 79 steps.

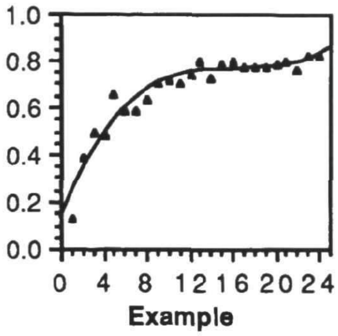

Figure 7: CELIA: Average results of exposure to sequence of examples. Average accuracy for reasoner predicting observed expert problem solving actions.

the student will have opportunities to use the indices and censors learned late in the test period.[2]

The approach has several advantages over previous approaches. The use of instruction through worked out examples, in conjunction with the student making predictions, shortens the feedback cycle, making learning more likely and more efficient. The retrieval and comparison of an additional case helps limit efforts at explanation. Third, the use of empirical adjustment of feature importance allows the student to become better at retrieving appropriate case pieces even when he is unable to explain failures.

Conclusions

In order for a case-based reasoner to improve its performance, it must acquire more cases and improve its case retrieval so that the right cases are retrieved at the the right times. When the reasoner uses part of a previous case to suggest an incorrect action, that case snippet needs to be marked as to what makes it not relevant in the current situation. If a part of a case would have suggested the correct action, that case snippet should be marked with the features of the current situation that make it relevant. The student must try to form an explanation based on the incomplete knowledge that he has.

Our approach to learning indices and censors offers the benefit of increased efficiency of learning through initial similarity based comparison to other cases and through immediate direct feedback available through observation of expert problem solving. Our case representation facilitates the learning process. Each snippet contains the pursuit of one primitive goal and the context in which it occurred. This enables the necessary comparisons between correct and incorrect contexts. The general approach allows the student to obtain the correct action for the current situation from the instructor, and attempts to distinguish when the correct and predicted actions are appropriate or not.

Learning indices, learning censors, and adjusting feature salience result in improving case retrieval. All three of these types of learning fit neatly within our general approach of learning by observing expert problem solving. Much of the power of CBR comes from retrieving useful cases, so this learning improves the case-based reasoner. Indices help the reasoner retrieve case snippets when they would be particularly relevant. Censors help the reasoner avoid being misguided by a snippet which may seem to be relevant but isn't. Improved knowledge of feature salience will lead to less consideration of superficial and spurious features during similarity measurement prior to use of indices and censors.

[2]It should be noted that some orders of presentation of examples will facilitate index learning. This ablation study did not take that into account. If example order can be controlled then learning can have a greater effect.

References

D. W. Aha. Incremental, instance-based learning of independent and graded concept descriptions. In *Proceedings of the Sixth Annual International Workshop on Machine Learning*, San Mateo, CA, 1989. Morgan Kaufmann.

R. Bareiss. Exemplar-based knowledge acquisition: a unified approach to concept representation, classification, and learning. Academic Press, New York, NY, 1989.

R. Barletta and W. Mark. Explanation-based indexing of cases. In *Proceedings of a Workshop on Case-Based Reasoning*, San Mateo, CA, 1988. Morgan Kaufmann.

K. J. Hammond. Learning to anticipate and avoid planning problems through the explanation of failures. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-86)*, San Mateo, CA, 1986. Morgan Kaufmann.

P. Koton. *Using experience in learning and problem solving*. PhD thesis, Massachussetts Institute of Technology, Cambridge, MA, 1988.

M. Lebowitz. Concept learning in a rich input domain: Generalization-based memory. In R. Michalski, J. Carbonell, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach, Volume II*. Morgan Kaufmann, Los Altos, CA, 1986.

R. M. Nosofsky. Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13:87–108, 1987.

M. J. Ratterman and D. Gentner. Analogy and similarity: Determinants of accessibility and inferential soundness. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ, 1987. Lawrence Erlbaum Associates.

M. Redmond. Combining explanation types for learning by understanding instructional examples. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, Hillsdale, NJ, 1989. Lawrence Erlbaum Associates.

M. Redmond. Learning from others' experience: creating cases from examples. In *Proceedings of the Second Workshop on Case-Based Reasoning*, San Mateo, CA, 1989. Morgan Kaufmann.

M. Redmond. Distributed cases for case-based reasoning; facilitating use of multiple cases. In *Proceedings of the National Conference on Artificial Intelligence (AAAI-90)*, Boston, MA, 1990. Morgan Kaufmann.

B. H. Ross. This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13:371–416, 1987.

S. Salzberg. Exemplar-based learning: Theory and implementation. Technical Report TR-10-88, Harvard University, Center for Research in Computing Technology, 1988.

R. L. Jr. Simpson. *A Computer Model of Case-Based Reasoning in Problem Solving*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, 1985.