An Empirical Evaluation of Model-Based Case Matching and Adaptation[1]

L. Karl Branting and **John D. Hastings**
Department of Computer Science
University of Wyoming
Laramie, Wyoming 82071-3682
{karl,hastings}@eolus.uwyo.edu

Abstract

Rangeland ecosystems typify physical systems having an incomplete causal theory. This paper describes CARMA, a system for rangeland pest management advising that uses *model-based matching and adaptation* to integrate case-based reasoning with model-based reasoning for prediction in rangeland ecosystems. An ablation study showed that removing any part of the CARMA's model-based knowledge dramatically degraded CARMA's predictive accuracy. By contrast, any of several prototypical cases could be substituted for CARMA's full case library without significantly degrading performance. This indicates that the completeness of the model-based knowledge used for matching and adaptation is more important to CARMA's performance than the coverage of the case library.

Introduction

One of the most striking characteristics of human problem-solving behavior is the ability to exploit multiple knowledge sources and reasoning techniques. This ability is important because most human problem solving occurs in an environment of incomplete knowledge. Automating this ability requires techniques for integrating multiple problem-solving paradigms in a flexible manner.

This paper describes an approach to integrating case-based reasoning with model-based reasoning for the task of predicting the behavior of complex and poorly understood physical systems. Under this approach, which is called *model-based matching and adaptation*, an incomplete causal model is used in assessing case similarity and adapting behavioral predictions in light of case differences. The next section describes how the task of predicting the behavior of a complex and poorly understood biological system arises from the overall task of rangeland pest management. The third section argues that a causal model that is insufficient to permit precise simulation may nevertheless assist case-based reasoning. The fourth section describes an implementation of model-based matching and adaptation in CARMA, a system for rangeland pest management. The fifth section sets forth an ablation study showing that each of the CARMA's model-based components makes a significant contribution to CARMA's predictive ability.

The Rangeland Pest Management Task

Our interest in integrating CBR with other reasoning paradigms arose from a project to develop an advice system for ranchers on the management of rangeland grasshopper infestations. Forage losses from grasshopper infestations have a significant economic impact, particularly in the mountain west. While grasshopper infestations can be treated with pesticides, the benefits of pesticide application are often outweighed by its costs, particularly when loss of beneficial insects and other environmental damage is considered. Providing advice about such infestations is a complex task because rangeland ecology is poorly understood and very complex. However, entomologists experienced in rangeland management routinely provide accurate advice to ranchers.

To clarify the problem-solving steps in the rangeland pest management advising (RPMA) task, we performed a protocol analysis of problem solving by a professor and several graduate students of entomology at the University of Wyoming experienced at this task. This protocol analysis indicated that entomologist problem solving consists of the following stages:

1. Determine whether grasshopper consumption will lead to competition with livestock for available forage.
   (a) Estimate the proportion of available forage that will be consumed by grasshoppers.
      i. Infer relevant characteristics of the grasshopper infestation, such as grasshopper species, developmental stage, and density, from observables.

[1] We wish to acknowledge the invaluable assistance of Professor Jeff Lockwood in the development of CARMA and of Scott Schell in providing a set of test cases. This research was supported in part by a grant from the University of Wyoming College of Agriculture.

ii. Determine the prototypical infestation case that most closely matches the current case.

iii. Adapt the consumption estimate predicted by the prototypical case based on the featural differences between the prototypical and current cases.

(b) Compare the grasshopper consumption with the proportion of available forage needed by livestock.

2. If there will be competition, determine what possible treatment options are excluded by the conditions of the case, *e.g.,* rainy conditions preclude the use of carbaryl bait, environmental sensitivity precludes all nonbiological treatments.
3. Estimate the expected economic costs/benefits of each acceptable treatment option by determining the proportion of grasshoppers that would be prevented from further forage consumption and egg laying under each treatment option, estimating the decreased probability of infestation in subsequent years if a given proportion of grasshoppers were prevented from laying eggs, and calculating whether the expected value of the benefits of each treatment outweighs its costs.

We have implemented this problem-solving process in a system termed CARMA (CAse-based Range Management Adviser). This paper focuses on the components of CARMA that perform steps 1(a)ii and 1(a)iii, estimation of the proportion of available forage that will be consumed by grasshoppers. Making this estimation requires predicting the behavior of a rangeland ecosystem, a physical system with an incomplete causal model.

Behavioral Prediction with an Incomplete Causal Model

A causal model for the behavior of a physical system is a model of the interactions among the components of the system that is capable of predicting or explaining the system's behavior. While many of the components of a rangeland ecosystem are known, the interactions among these components are only partially understood.

Figure 1 sets forth the most important of the qualitative causal constraints that influence forage consumption. Other information available for modeling rangeland ecosystems includes the following:

- The developmental stages of grasshoppers, including
  - The average length of each stage.
  - The proportion of lifetime consumption that occurs at each stage.
  - The average attrition rate at each stage.
- Some species of grasshoppers, termed *overwintering*, hatch late in the growing season, hibernate during the winter, and complete their development during the following growing season. Others, termed *nonoverwintering* species, hatch, lay eggs and die within a single growing season.

Figure 1: Qualitative relations in rangeland ecosystems.

- The significant production of forage at a location occurs during a specific portion of the growing season, termed the *critical forage growing period*, for that location.

Attempts have been made to construct large-scale numerical simulation models for grassland ecosystems (Rodell 1978). While such models often provide insight into the dynamics of ecosystems (Fedra 1991), there is a general recognition that the interactions affecting grasshopper population dynamics are too poorly understood and too complex to permit precise prediction through numerical simulation (Lockwood & Lockwood 1991; Pimm 1991; Allen & Hoekstra 1992).

As mentioned above, our protocol analysis indicated that part of entomologists' process of predicting forage consumption consists of comparing a new case to prototypical infestation cases. This suggests that entomologists use a case-based approach to this task. However, entomologists are capable of generating detailed causal explanations for their predictions, *e.g.,* in terms of the qualitative constraints shown in Figure 1. Part of this ability may come from associating stereotypical explanations with prototypical cases. However, entomologists can easily generate causal explanations of the effects of incremental variations on case facts. This strongly suggests that entomologists use the incomplete causal model to adapt and modify predictions associated with prototypical cases.

In summary, the available causal model of rangeland population dynamics is insufficient to permit precise numerical simulation. However, this causal model appears to play an important role in entomologists' problem solving, both in explanation and in adaptation of the predictions associated with prototypical cases. Automating entomologists' problem-solving ability therefore requires a computational mechanism for integrating this incomplete causal model with case-based reasoning.

Using Model-Based Reasoning To Assist Case-Based Reasoning

CARMA is a system for advising ranchers about rangeland grasshopper infestations. This section briefly overviews CARMA's case-based reasoning component and describes four different ways in which CARMA uses a causal model to assist case-based reasoning: case factoring; temporal projection; featural adaptation; and critical period adjustment.

Prototypical Infestation Cases

Since our protocol analysis indicated that entomologists estimate forage consumption by comparing new cases to prototypical infestation scenarios, we elicited a set of prototypical cases from an entomologist experienced at the RPMA task. These prototypical cases differ significantly from conventional cases. First, the prototypical cases are not expressed in terms of observable features (*e.g.,* "Whenever I take a step, I see 4 or 5 grasshoppers with brightly colored wings fly"), but rather in terms of abstract derived features (*e.g.,* "A low or moderate density of postwintering grasshoppers in the pre-adult stage"). Second, since entomologists see a very large number of specific cases, the prototypes represent generalizations of multiple cases, rather than specific cases. As a result, the feature values of these cases are often ranges rather than specific values (*e.g.,* low to moderate grasshopper densities and normal to hot temperatures). Finally, the prototypical cases are extended in time, representing the history of a particular grasshopper population over its lifespan.

CARMA's case library currently consists of eight prototypical cases. Each prototypical case is represented by a "snapshot" at a particular, representative point in time selected by the entomologist. In general, this representative time is late enough in the season that it is possible to determine the extent of the infestation with some certainty, but early enough that pesticide application is still feasible. An example prototypical case appears as Case8 in Table 1.

Case-Based Prediction of Forage Consumption

CARMA begins the process of predicting forage consumption by prompting the user for the observable features of the infestation. A series of rules are used to infer the relevant case features, such as the species, density, and developmental stage of the grasshoppers.

Factoring Cases Into Subcases A tract of rangeland often contains multiple, distinct grasshopper populations composed of species whose consumption characteristics vary greatly. Specifically, overwintering grasshoppers, which divide their consumption between two growing seasons, consume far less during the critical forage growing season than nonoverwintering grasshoppers. CARMA therefore factors the overall population of a case into subcases according to wintering types.

CARMA uses a model of grasshopper developmental stages to estimate the probable hatch and death dates of each grasshopper population given the current developmental stage, growing season dates for the location, and current date. If the hatch date occurs before the current growing season or the death date occurs after the current season, CARMA concludes that the grasshopper population is overwintering. Otherwise, the population is determined to be nonoverwintering. For example, the new case set forth in Table 1 is split into two subcases, SubcaseA and SubcaseB, based on wintering types.

Temporal Projection To predict the forage loss of a subcase, CARMA first retrieves all prototypical cases whose wintering type matches that of the subcase. Since prototypical cases are extended in time but are represented at a particular time, matching requires temporally projecting the prototypical cases forwards or backwards to match the average developmental stage of the new subcase. This requires simulating grasshopper attrition, which depends on developmental stage, precipitation, and developmental speed, which in turn depends on temperature, throughout the interval of the projection. An example appears in Figure 2.

The projected prototypical case whose weighted featural difference from the new case is least is selected as the best match. Feature weights are set using a hill-climbing algorithm that optimizes CARMA's predictive accuracy on training instances. For example, the prototypical case that best matches SubcaseA after projection is Case8, as shown in Table 1. Because the developmental stage of Case8 before projection is later than that of SubcaseA, Case8 must be projected backwards in time, causing grasshoppers that had been lost to attrition to be added back to the population.

Temporal projection aligns developmental stages but not necessarily dates. For example, the date of Case8 after projection is later than the date of SubcaseA because the hatch date of Case8 was later than that of

	Case8	New case		Case8 after projection
		SubcaseA	SubcaseB	
Wintering type	prewintering	prewintering	nonoverwintering	prewintering
Feeding types	grass 100%	grass 40% mixed 60%	grass 100%	grass 100%
Average stage	2.0	1.2	7.0	1.2
Density	(mod) = (10 14)	13.0	7.0	(11.2 15.6)
Date	September 8	August 20		September 2
Precipitation	(normal)	dry		(normal)
Temperatures	(normal)	cool		(normal)
Infest. history	(low mod-low mod)	mod		(low mod-low mod)
Range value	(low mod-low mod)	high-mod		(low mod-low mod)
Forage loss	(low mod-low)	?		(low mod-low)

Table 1: Case examples.

Figure 2: Projection of a prototypical case from PC to PC$'$ to align its developmental stage with new case NC.

SubcaseA. As a result, the developmental stage of the grasshoppers in SubcaseA on August 20 are the same as those of Case8 two weeks later on September 2.

Featural Adaptation The consumption predicted by the best matching prototypical case is modified to account for any featural differences between it and the subcase. This adaptation is based on the influence of each feature on consumption as represented by global feature weights. For example, a lower temperature value means lower forage losses, because lower temperatures tend to slow developmental speed, increasing grasshopper attrition. Thus, the forage loss estimate predicted by Case8—(low mod-low)—must be adapted downward somewhat to account for the fact that temperatures in SubcaseA (cool) are lower than in Case8 (normal).

The feature weights used in featural adaptation constitute a linear approximation of the function from derived case features to consumption amounts in the neighborhood of each prototypical case.

Critical Period Adjustment Consumption is only damaging if it occurs during the critical forage growing period of a rangeland. The forage loss predicted by a prototypical case must be modified if the proportion of the lifespan of the grasshoppers overlapping the critical period differs significantly in the new case from the proportion in the prototypical case. This process, termed *critical period adjustment*, requires determining the developmental stages of the new and prototypical cases that fall within the critical period and the proportion of lifetime consumption occurring in these developmental stages.

An example of critical period adjustment appears in Figure 3. Because grasshopper development in SubcaseA is ahead of that in Case8 (SubcaseA's developmental stage on August 20 corresponds to Case8's developmental stage on September 2), CARMA determines that SubcaseA applies to more of the critical period than Case8 because it will reach stage 5 by the end of the critical period, while Case8 will only reach stage 2. CARMA uses a model of grasshopper's rate of consumption at each developmental stage to calculate the proportion of lifetime consumption occurring before the end of the critical period. For example, 47% of SubcaseA's consumption occurs during the critical period, whereas only 6% of Case8's consumption occurs within this period. CARMA therefore scales the initial consumption estimate by (47 / 6) = 7.8.

After adaptation, the consumption predictions for each subcase (*i.e.,* behaviorally distinct population of grasshoppers) are summed to produce an overall consumption estimate. In the given case, the sum of predicted consumption of the two subcases is `high-moderate`.

In summary, CARMA uses model-based reasoning in four different ways to assist case-based reasoning. First, a model of grasshopper developmental stages is used to estimate probable hatch and death dates in

Figure 3: Critical period adjustment from Case8 to SubcaseA.

order to factor the new case into behaviorally distinct subcases. Second, a model of grasshopper attrition is used in temporal projection to simulate the attrition that would have occurred during the interval between the date of the new case and the representative date of the prototypical case. Third, featural adaptation constitutes a linear approximation of the function from derived case features to consumption amounts in the neighborhood of each prototypical case. Finally, critical period adjustment modifies the prediction estimate to take account of any difference in overlap between grasshopper lifespans and the critical forage growing season.

Evaluation

We performed an evaluation to determine the relative contribution of empirical (*i.e.,* case-based) and model-based knowledge to the performance of CARMA's consumption prediction module. To determine the contribution of model-based knowledge, we performed an ablation study in which the performance of the full CARMA consumption prediction module was compared to the performance of CARMA with various model-based components removed and to two different inductive methods: ID3 (Quinlan 1986) and least-squares linear approximation. To determine the contribution of empirical knowledge, the performance of CARMA was tested with the full multiple case library replaced with single case libraries.

The experimental case sets included ProtoL, Set1, and Set2. ProtoL is a library of eight prototypical cases generated by an entomologist, consisting of two overwintering cases and six nonoverwintering cases. Set1 consists of 15 cases generated by the same entomologist. Set2 contains 48 test cases whose features were randomly generated and whose forage loss predictions were estimated by a second entomologist.

Experimental Design

CARMA was tested using ProtoL as its case library. CARMA's global feature weights, used both in case matching and in adaptation, were tuned using a hill-climbing algorithm to optimize CARMA's overall predictive accuracy on Set1. The ablated versions of CARMA used the same global feature weights and case library as the full system.

ID3 was given ProtoL and Set1 as training instances. Because prototypical case features contain ranges, cases in ProtoL were treated as pairs of cases with the feature values associated with the minimum and maximum forage loss predictions of each prototypical case.

The linear approximation approach consists of a linear equation for forage consumption as a function of case feature values represented in quantitative terms. The coefficients of this equation were found using QR factorization (Hager 1988) to find a least-squares fit to the feature values and associated predictions of the cases in Set1 and ProtoL. As with ID3, ProtoL cases were treated as consisting of two cases corresponding to the minimum and maximum forage loss predictions of each prototypical case.

The single case library approach was evaluated by testing CARMA with the full case library replaced by individual members of ProtoL. These tests used the same global feature weights as the full system.

The accuracy of each approach was tested by comparing its forage loss prediction for each case in Set2 with the prediction of the expert. The qualitative difference between two forage loss predictions was calculated as the number of categories by which the predictions differ in the ordered set {`low, low-moderate, moderate, high-moderate, and high`}, so that `low` differs from `high` by four categories, the maximum possible qualitative difference. The results, which appear in Table 2, include the total qualitative error (qualitative difference between the prediction of the approach and the expert over all the test cases), the total number of incorrect predictions, and the average qualitative error per test case.

Because the maximum possible qualitative error is four (*e.g.,* `low` instead of `high`), a constant prediction of `moderate` consumption could never be off by more than two qualitative categories. A constant prediction of `moderate` is included for purposes of comparison.

Discussion

CARMA's average qualitative error was .42, as compared to an average error of 1.67 for a constant prediction of moderate consumption.[1] Removal of model-based knowledge significantly degraded CARMA's performance. CARMA's error rate was almost doubled by removal of featural adaptation (.79), removal of projection and critical period adjustment (.83), or by removal of all three (.85). CARMA was not tested with case factoring disabled. However, ID3's performance on unfactored cases, 1.00, was lower than CARMA's

[1] This error for a constant prediction is higher than the value expected if the expert predictions were uniformly distributed across the five qualitative categories. However, the expert predictions were skewed towards high and low predictions.

	Total error (qualitative)	Total number incorrect	Average error (qualitative)
CARMA	20	15	.42
Constant pred.	80	46	1.67
Ablation of Model-Based Knowledge			
CARMA w/o FA	38	23	.79
CARMA w/o CPA,P	40	24	.83
CARMA w/o FA,CPA,P	41	28	.85
ID3	48	28	1.00
Linear approx.	57	23	1.15
Reduced Case Library			
Case1	39	27	.81
Case2	65	26	1.35
Case3	23	17	.48
Case4	50	29	1.04
Case5	77	25	1.60
Case6	23	15	.48
Case7	25	18	.52
Case8	24	15	.50
Average	40.8	21.5	.85

Table 2: Summary of test results. P, FA, and CPA represent projection, featural adaptation, and critical period adjustment, respectively.

performance with all model-based reasoning other than case factoring disabled, suggesting that case-factoring is also an important requirement for performance in this domain.

Featural adaptation assumes that the function for forage consumption can be approximated by a linear equation in the neighborhood of prototypical cases. Given the large contribution of featural adaptation to CARMA's performance, it seems reasonable to wonder whether the forage consumption function can be globally approximated by a linear equation. However, the performance of the linear approximation (1.15) indicates that a linear function for consumption as a function of case features is a poor predictor.

The results of the reduced case library tests surprisingly indicate that a CARMA case library consisting of only Case6 or Case8 produced as many correct predictions as a library containing the multiple prototypical case set ProtoL, although the average qualitative error is slightly higher. One contributor to this high performance may be that the prototypical cases have ranges of values for many features and therefore effectively represent multiple cases. However, this result indicates that the completeness of the model-based knowledge used for matching and adaptation is more important to CARMA's performance than the coverage of the case library.

The effectiveness of model-based adaptation is illustrated particularly vividly by the fact that a case library consisting solely of Case8 performed nearly as well as the full case library. This result was unexpected because Case8 is an overwintering prototypical case with low predicted consumption (because most of the lifespan of the grasshoppers occurs outside of the critical forage growth period), while most of the cases in Set2 contain nonoverwintering grasshoppers with much higher predicted consumption.

Future Work

The most important weakness of the current implementation of CARMA's forage consumption prediction module is that it uses a single set of global feature weights for matching and for featural adaptation. Even if the consumption function can be approximated by a linear function in the neighborhood of prototypical cases, as assumed in featural adaptation, it doesn't follow that the same linear function is appropriate for all prototypical cases. Indeed, the observed poor performance of global linear approximation, shown in Table 2, suggests that linear approximations, and therefore feature weights, should be specific to individual prototypical cases. Moreover, while it is plausible that feature weights for matching should be the same as feature weights for adaptation, this hypothesis has not been tested. Thus, an important piece of future work is to test case-specific adaptation and matching feature weights.

An important limitation of the evaluation reported here is that the consumption predictions associated with Set2 were produced by a different entomologist than the entomologist from whom the prototypical cases were elicited. As a result, there may be inconsistencies between the testing set and the library of prototypical cases. We have subsequently obtained a set of 32 test cases produced by USDA entomologists experienced in this task and by the entomologist who was the source of the prototypical cases. This should permit a more reliable evaluation of CARMA's consumption estimation module.

Related Work

Several previous research projects have investigated the benefits of integrating case-based reasoning with model-based reasoning. However, these projects have generally assumed the existence of a correct and complete causal model. For example, CASEY (Koton 1988) performed diagnosis using model-based reasoning to assist both case matching and case adaptation. However, CASEY presupposed both the existence of a complete causal theory of heart disease and complete explanations of each case in terms of that theory. Because the causal model in CARMA's domain is insufficient for accurate prediction and the causal explanations associated with cases are incomplete, the assumptions underlying CASEY's matching and adaptation strategies are inapplicable to CARMA's domain.

Rajamoney and Lee (Rajamoney & Lee 1991) used a different approach to integrating case-based reasoning with model-based reasoning termed *prototype-based reasoning*. This approach uses a library of prototypes to decompose problems into familiar subproblems. Model-based reasoning is applied to the subproblems, a consistent composition of the subproblems is determined, and model-based reasoning is applied to determine the behavior of the resulting simplified model. As with CASEY, this approach presupposes a complete and correct (though not necessarily tractable) causal model. Similarly, Goel and Chandrasekaran's approach of using device models to adapt design cases presupposes that the device models are complete and correct (Goel & Chandrasekaran 1989).

Feret and Glascow (Feret & Glascow 1993) describe an alterative approach under which model-based reasoning is used for "structural isolation", *i.e.*, identification of the structural components of device that probably give rise to the symptoms of a fault. Cases are indexed by these tentative diagnoses, which are then refined using case-based reasoning. This approach, while appropriate for diagnosis, is ill-suited for behavioral prediction in the absence of faults.

Conclusion

This paper has described CARMA, a system that integrates case-based reasoning with model-based reasoning for prediction in rangeland ecosystems. CARMA uses four different mechanisms for model-based matching and adaptation: case factoring; temporal projection; featural adaptation; and critical period adjustment. An empirical evaluation showed that removing any of the latter three model-based mechanisms dramatically degraded CARMA's predictive accuracy. Moreover, CARMA performed significantly better with temporal projection, featural adaptation, and critical period adjustment removed than inductive approaches (ID3 and linear approximation) trained on unfactored cases, strongly suggesting that case factoring is also an important contributor to CARMA's performance.

However, any of several prototypical cases could be substituted for CARMA's full case library without significantly degrading performance. This indicates that the completeness of the model-based knowledge used for matching and adaptation is more important to CARMA's performance than the coverage of the case library.

References

Allen, T. F. H., and Hoekstra, T. W. 1992. *Toward a Unified Ecology.* New York, NY: Columbia University Press.

Fedra, K. 1991. Interactive modeling of environmental impacts. In *Proceedings of Automatic Control: Triennial World Congress*, 521–527.

Feret, M. P., and Glascow, J. I. 1993. Hybrid case-based reasoning for the diagnosis of complex devices. In *Proceedings of Eleventh National Conference on Artificial Intelligence*, 168–175. Washington, D.C.: AAAI Press/MIT Press.

Goel, A., and Chandrasekaran, B. 1989. Use of device models in adaptation of design cases. In *Proceedings of the Second DARPA Case-Based Reasoning Workshop.* Morgan Kaufmann.

Hager, W. 1988. *Applied Numerical Linear Algebra.* Prentice Hall.

Koton, P. 1988. *Using Experience in Learning and Problem Solving.* Ph.D. Dissertation, Massachusetts Institute of Technology. Department of Electrical Engineering and Computer Science.

Lockwood, J., and Lockwood, D. 1991. Rangeland grasshopper (orthoptera: Acrididae) population dynamics: Insights form catastrophe theory. *Entomological Society of America* 970–980.

Pimm, S. L. 1991. *The Balance of Nature: Ecological Issues in the Conservation of Species and Communities.* Chicago: University of Chicago Press.

Quinlan, J. R. 1986. Induction of decision trees. *Machine Learning* 1:81–106.

Rajamoney, S., and Lee, H. 1991. Prototype-based reasoning: An integrated approach to solving large novel problems. In *Proceedings of Ninth National Conference on Artificial Intelligence.* Anaheim: AAAI Press/MIT Press.

Rodell, C. F. 1978. Simulation of grasshopper populations in a grassland ecosystem. In Innis, G. S., ed., *Grassland Simulation Model.* New York, NY: Springer-Verlag. 127–154.